80.8
F2

OXFORD ECONOMIC PAPERS

(NEW SERIES)

CURRENT SERIAL RECORD
NOV 30 1953
U.S. DEPARTMENT OF AGRICULTURE

VOLUME 5 OCTOBER 1953 NUMBER 3

Editorial Board: T. WILSON (*General Editor*), J. R. HICKS, A. BRIGGS, C. M. KENNEDY, G. B. RICHARDSON (*Secretary*)

CONTENTS

The Full-employment Economy, with Special Reference to Wages Policy. By SIR DOUGLAS COPLAND 221

How Long does a Company Live? By GUNNAR LINDGREN . . 235

Marginal Preference and the Theory of Welfare. By JEROME ROTHENBERG 248

A Reply. By W. E. ARMSTRONG 264

The Effect of Taxation on Risk-taking. By PAUL STREETEN . . 271

Scarcity, Marxism, and Gosplan. By P. J. D. WILES 288

Competition and Control in the British Sugar Industry. By P. E. HART 317

OXFORD
AT THE CLARENDON PRESS
1953

OXFORD ECONOMIC PAPERS

(NEW SERIES)

OXFORD ECONOMIC PAPERS, of which eight numbers appeared at irregular intervals between 1938 and 1947, became a periodical in 1949 and is now published three times a year. The pre-paid subscription price is 25*s*., post free, and the price of a single copy 12*s*. net. Orders may be placed with any bookseller or sent direct to the Publisher.

THE JOURNAL consists, in the main, of articles by Oxford authors, but the editors will also be glad to consider contributions from elsewhere. As in the past, articles of considerable length, which embody the results of substantial research, are included from time to time. The editors also publish occasional articles on Economic History, Public Administration, and Scientific Method, provided these articles are likely to be of general interest to economists. Contributors receive 50 free offprints.

PAPERS should be sent to the General Editor, c/o The Clarendon Press, Oxford. All correspondence other than that dealing with contributions should be addressed to the Publisher:

GEOFFREY CUMBERLEGE, OXFORD UNIVERSITY PRESS
AMEN HOUSE, LONDON, E.C. 4

THE FULL-EMPLOYMENT ECONOMY, WITH SPECIAL REFERENCE TO WAGES POLICY

The Sidney Ball Lecture, Oxford, 10 March 1953

By SIR DOUGLAS COPLAND

THE topic to which it was suggested that I should address myself certainly falls within the defined scope of this Lecturership as dealing with 'modern social, economic and political questions'. It may be doubted whether any other topic could be more apposite at the moment than 'wages policy in a full-employment economy'. For this problem is not only a basic economic problem for all members of the free world with advanced industrial structures; it also raises political and social issues of the greatest significance and is as much a question for the political philosopher as it is for the economist. Indeed, in these days of possibly over-full employment, the techniques of the economist are not enough. We must all be social philosophers today if we are to offer solutions on basic economic policy that will endure in a society which, at one and the same time, places so much emphasis on social and economic security, and seeks to preserve the essential conditions of the democratic freedoms. I think Horace has it somewhere: 'I wish he were human and could mix virtue with wine.' This is a problem where it is necessary 'to mix virtue with wine', and therefore one of the most difficult of social problems, because the mass of people, however much command they have of virtue, find it a new and somewhat baffling experience to mix this with sudden and unaccustomed quantities of wine. That is precisely what has happened to many Western democracies in the past 6 years. The wine of full employment came suddenly and unexpectedly in rather greater doses than was anticipated, and it has imposed quite severe tests on social virtue.

1. Administrative Problems under Full Employment

In the Godkin Lectures at Harvard in 1945, *The Road to High Employment* (Harvard University Press), I made some observations on wages policy, among other basic economic issues that faced democracies under conditions of full employment. Though I approached the whole problem with more caution than some of the extreme advocates of full employment, the 'perfectionist school' with the motto enshrined on their banners—'more jobs than there are people to fill the jobs'—I find on reflection that I rather understated the administrative difficulties of evolving a wages policy that would relate wages to productivity, preserve reasonable margins for skill and other differences in conditions of work, avoid the ever-present tendency to income inflation, and promote mobility of labour. With Australian experience of collective bargaining in mind, it was

suggested that the periodic declaration by a central wage-fixing authority of a minimum wage adjusted to measurable changes in productivity, coupled with the automatic adjustment of that wage to changes in the cost of living, would ensure that the general level of wages would move in accordance with changes of prosperity. Moreover, the minimum wage would serve as the norm around which bargaining could take place in industries with a high and rising level of productivity, thus giving some incentive for mobility. For the rest, mobility could be encouraged by special aids to training, travel assistance, and differential rates for certain areas to which it was desired that labour should be attracted. This outline of wages policy was to be developed within a general framework of economic policy involving certain indirect controls over aggregate national expenditure, the distribution of the national income, public finance, and the foreign balance; all designed to sustain a high level of employment and to keep the economy in reasonable equilibrium. It was further assumed that such an economy would provide wide scope for private enterprise and retain the right of the individual to choose his avocation within a framework of broad social controls. Provided the level of employment was not extravagantly high and the wages tribunals operated as they did in the two decades before the war, the wages policy suggested might well have been appropriate to the new conditions. In the event, however, conditions of over-full employment operated in Australia and the country had 'to endure the rigours' of extravagantly high wool prices. At the same time, certain basic changes were made in the system of wage-fixing, so that the conditions of the post-war period were substantially different from those postulated in the Harvard lectures.

Meanwhile, other advanced social democracies, notably the United Kingdom, Sweden, and New Zealand, were facing similar situations with labour in short supply and a sellers' market both at home and abroad. Each in its own way was improvising a policy on wages and other basic factors in the economy as it sought to check the forces of inflation and to sustain productivity and enterprise. It was essentially a period of experiment in social policy in a situation that was unexpectedly favourable to sustaining an extremely high level of employment. At the same time, there was a rapid expansion of social services, increasing provision for economic security, a considerable growth in the percentage of national income going to labour or to the small farmer. The period was marked also by a hang-over of trade and currency restrictions from the war, reinforced by fresh restrictions designed to meet that post-war nightmare, the dollar problem. Thus the economies most under pressure to devise a new wages policy were working partially within a closed system and experiencing a degree of social change much more rapid than was generally realized. It

is little wonder that there was confusion of both thought and action, and that the advocates of full employment found themselves somewhat bewildered by the severity of the policy measures and the administrative task confronting governments, employers, and trade unions. One may express some surprise that the measures devised and the general atmosphere of tolerance contributed so much to hold a difficult situation in check. Now that all the countries concerned are facing rather less buoyant conditions it is an appropriate time to review the experience of the past 6 years, and to ascertain what it offers in either theory or practice towards the development of a wages policy in the new Welfare State with its goal of high employment, if not of full employment according to the perfectionist school.

So far nothing very positive has emerged from the rather scanty economic literature that has appeared on this problem, and one is perhaps entitled to observe that those who embraced the doctrine of full employment too readily are perhaps a little chastened by the turn of events. Well, I suppose that is the beginning of wisdom. I am assuming that the general objective is to retain the basic elements of a free society, to provide scope for the system of private enterprise to fill its accustomed role of sustaining a large measure of investment, to give the economy reasonable flexibility, and to operate it within a wider international economy. How far these basic assumptions would be accepted by economists as a whole is a matter for speculation. To those who are prepared to embark upon an increasing measure of direct control over labour and other resources, the problem of maintaining equilibrium and promoting an appropriate wages policy is a relatively simple matter provided the community as a whole is prepared to have its face lifted in the process. As such it is no longer an economic problem; the economist becomes a social philosopher and one who is not likely to be accepted readily by the free society. This is one of the lessons we learn from recent experience. The free society contains within it the elements of resistance to pervasive controls. It prefers freedom to order, some degree of instability to the deadening impact of regulation. One would expect employers and farmers to resist control, except perhaps where their own interests might be narrowly advanced, but an interesting reflection on the experience of recent years is that the trade unions as a whole resist any wages policy that impairs their freedom, their prestige, and their traditional role as guardians of the rights of their members. They have what might be called a sense of community which advocates of direct controls to reach their own given ends appear to lack. As a recent writer has put it: 'The welfare state is consistent with private economic power. The Socialist state is by definition opposed to it. It would convert private economic power back into public power.'[1]

[1] MacIver, *Democracy and the Economic Challenge*, p. 30. I have elaborated this theme in

2. Conditions of Equilibrium in the Free Society

This may seem to be far removed from the problem of wages policy in a full-employment economy, but it is necessary to be clear on the sort of society that one is postulating. I shall therefore consider the problem within the general social framework now outlined. In these conditions, wages policies would be developed as an integral part of a general economic policy designed to sustain a high level of employment whilst imparting equilibrium to the economy as a member of a wider international economy with which it has reasonably free intercourse in trade, finance, and currency. Moreover, the economy itself would be a blend of public and private enterprise providing scope for individual enterprise, yet with the maximum degree of social security, consistent with the maintenance of efficiency and the promotion of economic progress. In other words, it is the Welfare State as now developing in the Western world with its outposts in the United Kingdom, the Scandinavian countries, Holland, and the two democracies in the South Seas. In such an economy the conditions of equilibrium might be stated as follows:

1. That aggregate demand in terms of money does not exceed available supplies at the current prices, and that the level of prices is reasonably stable.
2. That the level of investment, public and private, is in balance with the level of domestic savings, plus or minus the external balance on capital movements.
3. That spending power is so distributed that the demands of each member of the community can be made effective on a reasonably equitable basis.
4. That the factors of production are sufficiently mobile to permit progress and change.
5. That the price structure is not distorted by the influence of monopoly.

If these conditions can be met, the economy will be in equilibrium, the path to economic progress will be assured, and the community will be able to enjoy the fruits of the Welfare State with a rising standard of living. It is not to be expected that 'smooth success will be strewn at our feet'. These are the conditions of long-term equilibrium. For a variety of reasons there will be fluctuations that will upset the balance in the short period, as there must be in any free economy. What is required is a self-adjusting mechanism that operates, within the framework of control, with increasing force the greater is the departure from equilibrium, as does the pendulum of a clock. This may be looking for perfection in a world of pressures and much short-sighted self-interest. It is well to recognize the simple

the Presidential Address to the Australian and New Zealand Association for the Advancement of Science, August 1952, under the title 'Authority and Control in a Free Society'.

truth that the body economic cannot be kept stable in the short period, and that any policy designed to correct at once all tendencies to instability requires the exercise of an intimacy of regulation and control that is inconsistent with the maintenance of the sort of society and economy I have postulated.

It is not possible in this lecture to state in detail the major measures of economic policy and control designed to achieve the general objective of long-term equilibrium. They are quite well established now, and the fact that they did not achieve equilibrium in any of the developing Welfare States need not deter us. They probably met with more success in the American and Canadian economies, which were rapidly expanding and permitted more scope for the free play of economic forces. That has its own lesson, and he who runs may read. But the circumstances of the United Kingdom, Australia, and some other countries were unusual in that the demand for labour was unexpectedly high, there was the great advance in social services and the total income enjoyed by labour, and the buoyancy of export income. These conditions are not likely to recur in the next few years at anything approaching the same tempo. So there is ground for assuming that, in a less hectic state of social change and economic expansion, the measures of control will be able to operate with reasonable hope of success. Moreover, there is some revision of attitude towards the controls themselves, and a revival of faith in such old-fashioned but wholesome measures as a two-way operating monetary policy, an increasing emphasis upon saving, and a willingness to contemplate a mild surplus of resources in the short period. This change is in part a conscious act of policy, in part an inevitable response to a world that is growing harsher, as some countries restore their shattered fortunes and others thrust out to new conquests. In these conditions, therefore, we may expect the now familiar weapons of monetary policy with a flexible rate of interest; of fiscal policy with a variable budget surplus or deficit; of investment policy with more careful attention to the balance between public and private investment appropriate to the nature of the economy; and of credit policy working within a more flexible system of interest rates to work much more effectively in promoting economic equilibrium than was possible in the immediate post-war years. I am not unmindful of the many hazards that face governments and economic administrators in applying these policies to conditions as they change. The techniques are by no means fully developed, and public opinion has yet much to digest before they will be generally accepted.

There is one additional hazard that has a direct bearing upon wages policy, which I am assuming will work within the framework of the controls already suggested. This is the variation in money income due to

changes in the terms of trade. This hazard is much more troublesome for countries like Australia, with raw materials and foodstuffs predominating in the export trade, than it is for advanced industrial countries with an export trade in finished or capital goods. In the former, the terms of trade are subject to violent short-term fluctuations as well as long-term shifts, and this can be very disturbing to the level of money income in the short term and to the stability of the economy. This is the post-war experience of Australia, where by far the most inflationary force was the great rise in exports, especially in wool, in 1949–50 and 1950–1. None of the policy measures suggested can cope with the impact of sudden changes in export income on the general level of money incomes; certainly it is impossible to meet the problem by wages policy alone. Thus, in Australia, export prices rose from 348 in 1948–9 (base 100 for average of 3 years ending June 1939) to 690 in 1950–1. The main force contributing to this was the price of wool, which in the same period rose from 365 to 999. The value of exports rose from £543 mn. in 1948–9 to £982 mn. in 1950–1, whereas in 1946–7 the comparable figure was only £309 mn. It is not surprising that there was a great increase in money incomes as a whole, and in costs. The Australian economy is highly responsive to sudden changes in the level of export income, and between 1948–9 and 1950–1 the estimated national income rose from £1,937 mn. to £3,101 mn. This is a disturbing force that cannot be wholly countered by the less direct controls, such as variations in the exchange rate or a differential export tax. The exchange adjustment would have fallen too severely on some exports and with insufficient severity on wool, and, to be effective, the differential tax would have involved what the export producers would regard as almost savage confiscation. The most satisfactory measure was a substantial levy on gross proceeds to be placed in a stabilization fund, but this too was resisted, and the only action finally taken was to impose a 20 per cent. levy as an advance payment of income-tax. In New Zealand more attention has been given to this problem, and measures have been devised through stabilization funds to soften the impact of variations in export prices on the level of money incomes and on the economy as a whole. Some such action is essential if an economy subject to sudden changes in the terms of trade is to be held in reasonable balance, and to avoid the devastating effect of rapid changes in the level of money incomes and costs.[1]

3. Money Wages and Productivity

This long excursion into the nature of the economy we have in mind, and the general framework of controls necessary to maintain long-term equi-

[1] For these problems as affecting Australia see my *Inflation and Expansion: Essays on the Australian Economy* (Cheshire, Melbourne, 1951), especially pp. 72–89.

librium and to ensure a self-adjusting mechanism that will operate to mitigate the tendency to short-period departures from equilibrium, is essential if a realistic approach is to be made to a wages policy. This policy cannot be considered in a vacuum in terms, either of the more specific economic issues involved, or of the broader issue raised by the social milieu. The first problem to be faced is whether a steady rise in money wages as a whole in response to rising productivity is to be contemplated, or whether money wages should be kept stable and the increased productivity be encouraged to express itself in a fall in the general price level. In some respects this is rather an academic question, because it may be doubted whether the general price level for the international economy as a whole can be brought under control. Secular movements will continue mainly under the influence of long-term changes in the prices of basic materials and foodstuffs that enter into international trade. If the balance of demand and supply went against the prices of foodstuffs and raw materials, the industrial countries would be unlikely to adopt a policy of exchange depreciation to sustain their price levels, though this device could be used by the producing countries to soften the impact of deflation. Much would depend upon the rate of price change. Marshall could argue that a period of expansion with steadily falling prices, as in the seventies and eighties of last century, brought the best of both worlds. It provided the necessary stimulus to enterprise, and automatically transferred to labour and fixed incomes an appropriate share of the increased productivity. Thus it provided an automatic rise in living standards. But the economic controversies of that idyllic period are not altogether reassuring on that point. They were the days of the bimetallic discussions, of Bryan's 'cross of gold' and the heresies associated with cheap silver. Today the social pressures against a relentless fall in prices would be overwhelming, and the trade unions and farmers' organizations, by far the most powerful groups in modern democracies, would combine for different reasons to resist a long-drawn-out decline in the general price level. Moreover, such a movement would raise awkward problems for all modern governments with their heavy obligations in fixed interest-bearing debt. So we may assume that on this first issue the desideratum of policy would be to register rising productivity in higher money wages and have a rising rather than a falling price level. Stability of the price level would be a permissible goal, but the long-term trend would be upward rather than downward. In so far as wages policy was a contributory factor to a moderate secular rise in the general price level, it would thus be consistent with the needs of an expanding economy, the pressures operating in the social framework of modern free communities, and the financial burdens placed upon modern governments.

4. Determination of a Minimum Wage

A wages policy assumes some norm around which actual earnings will fluctuate according to differences in skill and conditions of work, and differences in the relative prosperity of industries. The determination of this norm gives rise to problems on which there is a wide difference of opinion. In Australia, the decision was taken many years ago when the system of centralized wage-fixing tribunals was established, and over the past 30 years the minimum wage has become the norm. A similar situation exists in New Zealand, and any tendency to vary the minimum wage system would be resisted by a majority of the interests concerned, as well as by public opinion. Under the post-war conditions in Australia, this minimum wage has risen substantially, firstly because it is subject to automatic adjustment to the cost of living, and secondly because the principal tribunal (the Commonwealth Court of Conciliation and Arbitration) has granted increases in the basic minimum itself. Moreover, under the more prosperous conditions operating in all industries, especially in rural industries, the minimum wage has been extended to a larger number of workers, until it is now applicable to practically the whole of the work force. The basic additions to the wage have been made after review by the court of the prosperity of the economy as a whole. This review has now become a tedious affair at which evidence is tendered by the parties concerned or their expert witnesses, so that by the time a decision is made the circumstances may have changed. Thus in the last review in 1951, when the court added £1 per week to the wage, and raised the basic wage applicable to females from 54 per cent. of the male wage to 70 per cent., the decision was announced on the same day as the Treasurer introduced a budget designed to check inflation. But the system has the merit that it imposes upon some independent authority the task of determining a basic minimum wage that has some relevance to changes in productivity. Actual earnings tend to rise above the minimum wage in periods of active demand and to approximate thereto when demand recedes. By 1950–1 the increase in the money rate wage for men was 120 per cent. above the 1938–9 level, and for women 177 per cent. In real terms, when allowance is made for the cost of living, the increase for males was approximately 20 per cent. and for females 50 per cent. The greater rise for females is due to the levelling-up process when the female basic wage was raised to a higher proportion of the male wage. If in addition to this rise in the minimum wage, allowance is made for the higher level of actual earnings, for the increased provision of social services, for the indirect benefit of a high degree of economic security, and for the introduction in 1947 of a 40-hour week, the improvement in the standard of living as a whole of the lower paid workers is quite remarkable. This movement continued the

trend commenced during the war, when wages as a percentage of the net product in manufacturing rose from around 50 per cent. pre-war to 54 per cent. in the later years of the war and to 59 per cent. in 1949–50. Thus a large part of the improvement in the terms of trade and the increase in gross output due to full employment was transmitted to labour. But the preponderance of farm income in the years of rising prices brought down the percentage of wages and salaries to personal income as a whole. Before the war this percentage fluctuated around 55 to 59 per cent. It rose rapidly to 67 per cent. during the war. Then, in spite of the rising standards for labour, it fell to 54·5 per cent. in 1949–50, and in the year of the wool boom 1950–1 to 51 per cent., such was the impact of the great increase in personal income in the export sector. Had wages responded more rapidly to the rise in export income, the income inflation would have been greater and the general level of money costs higher. As wool prices receded by 40 to 50 per cent. in the following year, and the general level of export income fell, the higher level of wages and costs would have been embarrassing, though it might have been argued in 1950–1 that there was a case for a higher minimum wage and earnings at that time. But export income is exposed directly to the hazards of fluctuations in export prices, and any determination of a minimum wage as a norm in relation to productivity must take account of this fact. The fixed minimum would be expected to move more slowly than the level of export income, and in fact in 1951–2 there was a rise of some 15 per cent. in the basic minimum on account of continued increases in the cost of living, at a time when the export industries as a whole had experienced a reversal in their fortunes.

To Australia, with its long experience of fixing a national minimum, this movement in the basic wage was not altogether surprising in the circumstances. Does it offer any light on the problem of wages policy in a full employment economy? Apart from some considerations regarding differential rates to be discussed later, it is probably true to say that in the circumstances the minimum wage, and still more the actual rate of earnings, was about as high as the conditions would permit. Whether it can be sustained depends upon movements in the terms of trade, and on increases in general productivity in industry. The matter is still in doubt and is being argued before the Arbitration Court at the moment. In a more stable economy, where variations in the terms of trade and the value of export income are less spectacular, the fixation of a national minimum might be subject to fewer hazards. But this course of action is outside the normal experience of employers and trade unions, and is opposed on the grounds that it removes the determination of wages from the parties to the contract, and in particular is detrimental to the prestige of the unions and to the exercise of their normal functions. The Australian experience has little

relevance to conditions in, say, the United Kingdom or Sweden. Yet it has been found in practice that some central control over the general rate of wages is necessary if the improved position of labour under full employment is not to result in a rate of money earnings that unduly inflates the level of costs, and creates a wage rate that is not tenable when the economy becomes more exposed to the forces of world competition. The position has been met in part in the United Kingdom by the exercise of a policy of restraint on the part of the unions. This is to be preferred to any action directly by the government to impose a *maximum* wage, and, in the official view of the unions, to any move designed to establish a tribunal with authority to fix a basic or minimum wage. But the unions are themselves faced with a dilemma. If they exploit the increased power full employment gives to labour, they are likely to add to inflationary pressure, raise costs, and distort both the internal structure of industry and the balance of payments. In these conditions, controls and restrictions will have to be applied with increasing severity to investment, profits, prices, and imports, and it would be strange if money incomes and wages escaped the net of control. On the other hand, any central determination of a wage norm is an act of restraint on their part, unless it be imposed by the government, and any such action would be anathema to the unions.

5. Trade Unions and Wages Policy

These problems have recently been discussed by a group of Swedish writers from whom one gathers that, in spite of differences in procedure, the experience of Sweden in the post-war period does not differ so much from that of Australia.[1] There has been a levelling-up of the average wage for the lowest paid sector, a general pressure of rising wages on rising prices, a great improvement in the share of the product going to labour. Despite the increase in prices, the traditional pressures for higher wages have, in the short period, probably worked more favourably to labour than counter-measures designed to eliminate the inflationary pressure of rising wages. But this can only be the short-period result. Experience has now shown that the inflationary movement cannot be continuous. It affects the level of productivity, the rate of new investment, the foreign balance, the liquidity of the money market, and the rate of interest. Increasing restrictions, taxation policy, and controls can soften the impact of this on the economy, but sooner or later there is a decline in demand and a recession in economic activity that places an automatic check on the level of money incomes, even in an economy like the Australian, where so many money incomes are adjusted automatically to changes in the cost of living. One of the Swedish writers remarks on this short-term exercise in wage

[1] *Wages Policy Under Full Employment*, edited by Ralph Turvey.

bargaining that 'profit margins must be severely squeezed between the high wages which the trade unions can secure when there is full employment and a purchasing power restricted by taxation'.[1] Surely this shows a serious misconception of how the free economy works, and of the part played by profits. This view is resisted by Erik Lundberg, who rightly remarks: 'It is clear that an economic policy which aimed at compressing profit margins by allowing wage increases would be only transient.'[2] The main point about this argument is the assumption that the full-employment experience of the post-war period was to be perpetuated. Even in a closed economy it contained within it the seeds of its own destruction; in an open economy it is in the long run, and not the very long run, a *non sequitur*. So I see no way out of this dilemma for the unions, if they want to maintain reasonable stability of employment and activity, but to accept the principle of some central action to establish a minimum for the lowest paid sector and to exercise some restraint in bargaining for actual rates above this minimum. This approximates to the policy of the British unions, but it creates an awkward situation for the unions if the members press for what the traffic will bear in conditions of excess demand for labour. That position could be met by some bonus system on high profit levels earned through large turnovers, and not through price increases. Such a system would be far less inflationary than constant pressure for higher wage rates, and would provide a basis for a higher level of saving, and therefore of continuous investment, which is one of the major difficulties confronting the full-employment economy. But this procedure might also cut across trade-union policy, which has not been very ready to embrace any arrangement savouring of profit-sharing. One may well observe that the full-employment economy must involve some changes in the traditional outlook of the unions. After all, security of employment makes a big difference to the social structure in which the unions operate.

Some highly interesting and fruitful thought is being given to this problem in the Netherlands, where a new organization known as 'The Foundation of Labour' was set up after the war to deal with problems of industrial relations and to consider the general framework of economic policy in the immediate post-war conditions. This body is representative of employers and employed and has the official blessing of the government. It appears to have succeeded in holding wages at a predetermined level, avoiding industrial disturbances, imparting reasonable price stability to the economy, and promoting a steady rise in productivity. With the emergence of higher profits as recovery proceeded, the problem has now arisen whether wages should rise. The Secretary of the Netherlands Federation of Trade Unions, Mr. Vermeulen, has taken the view that no

[1] Gosta Rehn in Turvey, cited above, p. 49. [2] p. 58.

rise should take place in the money wage rate on this account and he seems to have persuaded his trade unions that there is a better alternative. He would impose a special percentage levy on profits and place the proceeds in a central investment bank which would be authorized to issue shares to workers. These shares would be negotiable for investment purposes, such as the purchase or construction of a house, the higher education of children, the payment of insurance premiums, &c. The benefit would thus go to labour in a form that would sustain and even increase investment and productivity without imposing undue strain on the foreign balance or the internal price structure. How far employers would be prepared to have profits taxed for such a purpose remains to be seen, but it is a stimulating proposal to come from the trade unions. The scheme is described in some detail by Mr. Vermeulen in an article in the forthcoming issue of *The International Labour Review*.

6. Differentials for Skill and among Industries

If some scheme for determining the wage rate and total remuneration of the unskilled worker could be put into operation on the lines I have suggested, it would enable the full-employment economy to avoid two other trends in the wages policy that in the long run are damaging to the efficiency and stability of the economy. The first of these relates to differentials for skill and conditions of work, and the second to differentials among industries according to efficiency, productivity, and profit-earning capacity. Both have a very important bearing on the mobility of labour. As regards the first, it is generally agreed that security of employment and excess demand for labour have encouraged a narrowing of the margin between skilled and unskilled wage rates and earnings. This has been very marked in Australia, where the average unskilled rate as a percentage of the average skilled rate has risen from approximately 75 per cent. pre-war to 85 per cent. On another test, namely, rates in lowest paid industry as a percentage of rates in highest paid industry, the movement has been from 73 to 80 per cent. Women's rates as a percentage of men's have risen from 55 to 67 per cent., and the latter figure does not fully reflect the influence of the recent decision of the court to raise the basic women's rate. This narrowing of differentials for skill is characteristic of a period of expansion, but if it is to become a permanent condition under full employment, it is clearly a matter on which the unions and the country as a whole must make some basic decision. In Australia the position has been complicated by a change in the Arbitration Act under which Conciliation Commissioners are appointed independent of the court, with the power to determine margins and other matters. This has not only separated the determination of margins from the determination of the basic wage; it has

produced the anomaly of two conciliation commissioners giving different decisions on margins in the same trade. One commissioner refused increased margins for metal trades workers covered by federal awards, giving as his reason that it would be conducive to inflation, while another a few weeks later granted an increased margin of 15*s*. per week for metal trades workers in the Australian Capital Territory. The decision of the first commissioner was fundamental and seems now to have been accepted by the unions, who are at the moment defending the basic wage and the 40-hour week in a case in the court. This levelling process is perhaps not as marked in other countries as in Australia, but it has been the general tendency under the impact of full employment. On the procedure I have suggested for dealing with the lowest rates, there would be ample scope for retaining appropriate margins for skill and giving the wage structure the flexibility required for incentive on the part of labour and for improvements in general productivity.

This preservation of margins for skill has also an important bearing on the mobility of labour. Differential rates in industries according to their relative efficiency and profit-earning capacity are even more important in promoting mobility of labour and imparting to the economy as a whole the flexibility necessary for industrial progress. Whilst industry generally would not be permitted to pay rates below the agreed minimum or the margins for skill, the existence of differential rates as between industries should not be discouraged, provided these differentials register differences in real efficiency, inventiveness, and the capacity of the industry to meet new and growing demands, and are not the result of monopolistic practices. In an economy where there is a considerable measure of insecurity for labour these differentials exist in the nature of normal bargaining. In a full-employment economy neither the employer nor the worker in the less efficient or the decadent industry feels the impact of competition, and the tendency is for wages in these industries to approximate to the rates current in the more progressive industries. It requires little imagination to realize how serious a brake such a situation will put on industrial progress. Full employment may well put a premium on inefficiency and backward techniques unless a conscious effort is made to ensure that the adventurous entrepreneur is able to command sufficient resources to embark upon his enterprise. This is as true of national works of an essential character as it is of private enterprise, and in practice it has been found necessary to offer some margins to attract labour to areas where important construction work is in progress. But if the basic norm is already high under the impact of a favourable bargaining position for labour, the differential rates may prove a handicap to the economic development of the project in question. So it is important in this matter, as in others, to

exercise restraint on the determination of an average rate for the lowest paid sector of labour.

This general outline of policy on wages is consistent with the operation of the broad economic controls I have postulated for the preservation of equilibrium in a full-employment economy of the type assumed. It is not the economy of the immediate post-war period, when demand was continuously in excess; it is an economy in which some conscious effort is required to sustain demand at the appropriate level, not all the time, but part of the time. In such an economy it is easier to adjust policy in the short period to the long-term needs, and thus to avoid extravagances that cannot but result in inefficiency and instability. In such an economy and the social structure it implies, wages policy is but one element of an economic policy that imposes new restraints and obligations on groups and individuals. In these circumstances, it cannot be assumed that traditional wages policy will be entirely appropriate, or that the functions and activities of trade unions will be unaffected.

HOW LONG DOES A COMPANY LIVE?[1]

By GUNNAR LINDGREN

CONTENTS

1. The Development of Limited Liability Companies 1913–42 . . 236
2. The Usefulness of the Material 238
3. The Companies' Activity and Size 240
4. The Formation of Companies 241
5. The Retirement of Companies 242
6. Old and Young Enterprises 243
7. The 'Death-risk' for Limited Liability Companies . . . 244
8. The Extreme Mortality of Young Enterprises . . . 245
9. Summary 247

One of the weightiest arguments in defence of the freely competitive economic system is that the pressure of competition forces enterprises to safeguard economic efficiency—at the risk of retrogression or complete annihilation. Under such conditions it would be of great interest if, with impressive figures, one could measure the power of the competitive pressure and its long-run influence on the development of efficiency.

From the nature of the problem, it is very difficult to make a statistical measurement of this process. One is forced to limit oneself to general indicators of various kinds. One such intrinsically useful indicator ought to exist in the 'turnover' of enterprises. If the process of thinning-out and regeneration is strong, one should expect a high frequency both of commencements and retirements, and this can reasonably be interpreted as active new enterprise initiative and a hard pressure of competition on the existing enterprises. Conversely, a low turnover of enterprises would indicate, fairly definitely, weak competitive pressure and difficulties in the way of new enterprise initiative.

Figures for enterprise formation and destruction at present exist only to a very limited extent in Sweden. The only regularly published data on enterprise formation are the reports compiled by the Patents and Registration Office and included each year in the *Statistisk Årsbok* (Statistical Yearbook), concerning newly established limited liability companies. No such general information exists on the mortality of enterprises. No doubt studies have been made, but they have included only limited sections of our economy. In such a case, the following account of the conditions in the Swedish limited liability company system may be of some interest.

[1] Translation of an article from the Swedish journal *Balans* 1949: 4 with the title 'hur länge lever ett aktiebolag?'

1. The Development of Limited Liability Companies, 1913–42

During the last decades the limited liability company form has, of course, greatly increased in significance. From before the First World War until 1942 the total number of such companies increased from about 7,000 to about 20,000—an increase, that is, of something like 180 per cent. It is difficult to find any natural comparison. If one chooses at random the population increase, one finds that, during the same period, there was an increase of 14 per cent. Possibly the development of real gross national income could be used as a measure of economic achievement. In round numbers, it increased by 125 per cent. between the same years. A third conceivable comparison can be made with employment in manufacturing, trade, and communications, which between the censuses of 1910 and 1940 grew by about 50 per cent.

The increased spreading of the company form is, of course, an indication of the continuing industrialization and commercialization of an economy —companies occurring most infrequently in the weakest sector of the industrial processes, agriculture and handicraft. To this it must be added that this form of enterprise has achieved such popularity that it is used even in smaller enterprises. The development has led to the situation where this form has become the dominating one for all enterprises above a certain relatively small size, outside agriculture. It may be calculated that at present about 50 per cent. of all employment in the private sector is given by limited liability companies.[1]

In Fig. 1 a more detailed picture of this development is given. The diagram shows the development, year by year, of the numbers of existing, newly formed, and retiring companies, based on material from the Patents and Registration Office. In order to give an impression of the proportions in the development, Fig. 2 shows the newly formed and retiring companies as a percentage of those companies which were in existence at the beginning of each year. Such a calculation probably does not mean very much for the newly formed companies, since the more natural comparison ought really to be with the amount of enterprise initiative, actual and potential, from which company formation derives. However, no such measure is available. For the retiring companies, the diagram's 'mortality rate' gives a better impression of the intensity than do the absolute numbers in Fig. 1.

The figures used in the diagram were derived in the following way. By the Company Law of 1910, companies became obliged to send copies of their balance sheets to the Patents and Registration Office. In a perusal of the returns for an entirely different reason from the present one, the

[1] The figure is derived from an investigation of enterprises made by the author, but as yet unpublished, using material from the obligatory accident insurance.

possibility of bringing together cases of liquidation and bankruptcy showed up quite incidentally, though without giving the distribution of the cases between the two. The rather considerable number of companies which

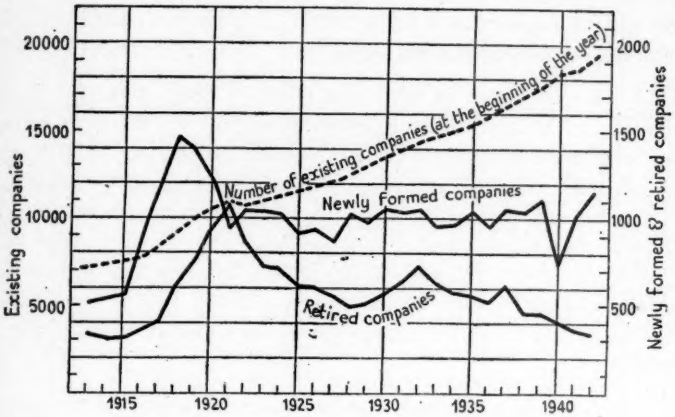

FIG. 1. Swedish limited liability companies 1913–42.

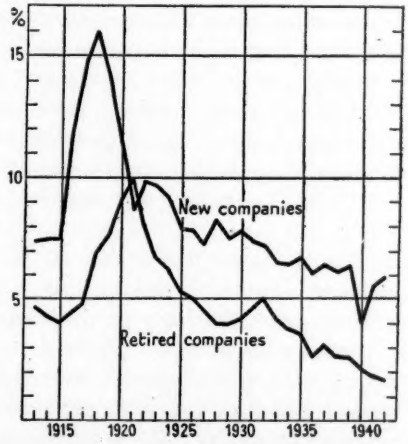

FIG. 2. Newly founded and retired companies as a percentage of companies existing at the beginning of the year.

ceased to forward returns was also considered as 'retired'. In such cases, the year of the company's retirement was assumed to be the calendar year after the last balance-sheet was received. This data was put together with the information made public on the number of newly formed companies,

corrected for such companies as expressly stated themselves to be not carrying on activity. Using this information about new formations and retirements, the development of the number of existing, actually operating companies could be calculated.

2. The Usefulness of the Material

If the material shown in the above diagram could be used, one would have a picture of the 'turnover' of enterprises in a very large sector of the Swedish economy. There arises the question, though, of whether the figures for new formations and retirements really reflect the actual changes. Are they, instead, possibly determined in the main by conditions which, from the real economic point of view, are arbitrary; by the flexible utilization, that is to say, of the possibilities for organization and reconstruction which the limited liability company form affords? In this connexion one can set up quite a few reservations against the material. Some of the most important are:

(1) The fact that a company continues to exist as a legal unit does not at all imply that one is dealing with one and the same economic enterprise at all times. As regards reconstruction, there may be such great changes in the real economic conditions within the company that it would be more correct to speak of new formation of an enterprise than of the reorganization of an existing company. An extreme case of this exists when a company is completely emptied of its real content and only the legal shell itself remains—the purpose, perhaps, being that the company may be transferred to a new owner who gives it a completely new real content. By this means a part both of new formations and cessations do not appear in the statistics.

(2) Company formation in many cases implies only that an enterprise, which has perhaps been in existence for a long time, is legally clothed in new garments. So far as this is the case, the diagram's curve of newly founded companies shows a more or less considerable unreliability as an indication of the actual formation of enterprises. Among the new formations there are surely also quite a lot of 'empty shells' which perhaps only after several years, sometimes perhaps never, will have any real content. In these cases the diagram registers the company formations too early. In other cases a newly formed company may be considered to be only the direct continuation of an old company which for some reason or other it has been desired to liquidate. In the diagram's curves one thus gets instances of both new formations and cessations where actually the company itself continues to exist as the same economic enterprise.

(3) That a company ceases to be registered does not necessarily mean that an economic enterprise has closed down, as has just been shown. But even where this is the case the conditions can vary considerably. Many cessations of enterprises are decided by factors quite other than the pressure of economic development. If, for example, a company is dissolved by reason of the owner's death, or because it was formed for a special task which is now not required, this is, from the point of view of economic development, something completely different from a dissolution deriving from a compulsory suspension of payments.

(4) Among the economically 'living' companies there are many with decidedly special functions, such as the administration of financial assets. It is conceivable that such companies are subject to completely different laws of development from the real 'producing' companies. If such companies occur quite frequently, and if their patterns of development are essentially different from that of the producing companies, then the diagram's curves for births and deaths in the sphere of limited liability companies do not give a good impression of average life conditions of companies in general.

It would not be difficult to extend this list of reservations which must be made in the economic interpretation of the diagram's curves. Unfortunately it has not been possible to undertake the reworking of the individual observations that these reservations would really require. In such circumstances one is inclined to ask whether the material, in broad outline, tells anything at all about the actual development of enterprise formation and retirement. Long familiarity with the material, however, has given the impression that one need not be too sceptical. On certain points, during the process of the work, it has been possible to estimate the size distribution of those factors discussed, no doubt with significant margins of error. The results indicate that the curves, in the main, bear the impression of the actual situation. To this it must be added that the various sources of error influence the figures in different directions so that the errors to a certain degree cancel out. Other errors appear to influence the level rather than the directions of movement of the curves, which implies that the material should at any rate be able to say something about the tendencies in the development. It is also worth noting that a theoretical analysis of the conditions which may be thought to influence the turnover of enterprises does not make the development shown in the diagram improbable. Finally, one can point to the general agreement of the results with those of the overseas (mainly American) investigations concerning similar problems. Therefore it does not appear incorrect—since a more thorough investigation is not available—to interpret the curves as, in the main, an expression of the actual conditions, i.e. that they should be

considered to give a certain, perhaps somewhat blurred, exposition of the order of size and the development of the 'turnover of enterprises' in the sphere of limited liability companies.

3. The Companies' Activity and Size

If one accepts this material as usable, the next question is: What sort of companies dominate the figures? This question involves partly the sphere in which the companies are active and partly the size of companies concerned.

Concerning the distribution of companies between different sectors of the economy, one may use the investigation which was undertaken in the preparation of the 1944 redrafting of Swedish Company Law. The investigation applied to the year 1934, but the results may be applicable to the whole period shown in the diagram. Of all companies, 35 per cent. were classified as 'manufacturing' companies. The next in order of size was 'wholesale and retail trade' with 31 per cent., and 'building activity and the administration of housing property' with 14 per cent. Together these three groups accounted for about 80 per cent. of the number of companies. If one adds 7 per cent. for 'trade in stocks and shares', and 5 per cent. for 'transport and communications', one may obtain a sufficiently diversified picture of the branches of the economy that dominate those statistics which use only the *number* of companies, without giving any more precise details on the size of capital or the number of employees.

It is not easy to decide on the measuring-rod to be used in determining the size of enterprise. If one uses the number of employees, it is mainly because this criterion is easy to understand. This criterion is also of special interest from the social viewpoint if one desires to obtain a picture of the significance of enterprise formation and cessation as a determining factor in the development of the labour market.

If one distributes the total number of companies existing in 1942 according to the size of the employment given, one obtains the results shown in Fig. 3, on p. 241.

The diagram gives clear evidence of the dominance of the larger companies as soon as one chooses another measure than the number of companies themselves. The differences are very marked between the smallest and the largest size-groups in the diagram. The smallest size-group contains almost 100 times as many companies as the largest, but only about a tenth of the latter's total number of employees. If one considers that companies with over 500 employees are 'large', then one finds that about 260 large units, a good 1 per cent. of the total, account for as much as 45 per cent. of the total employment given by companies. The medium-sized group, with 101 to 500 employees, which makes up almost 5 per cent.

of the number of companies, gives approximately 25 per cent. of the employment. The remaining 94 per cent. of smaller, or small, companies employs the remaining employees, less than one-third of the total.

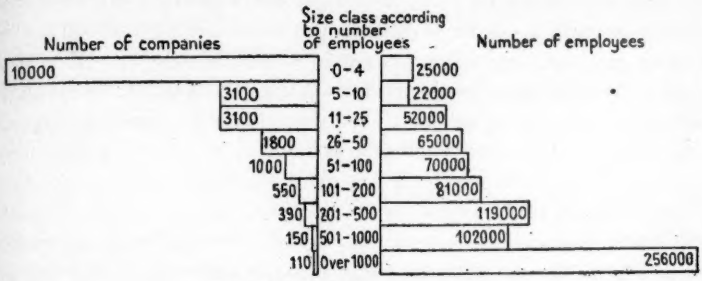

FIG. 3. Size of companies, 1942.

The figures used show that the data on the *number* of companies probably cannot help but be dominated by small companies. Not only existing but also newly formed and retiring companies may to a quite overwhelming degree consist of enterprises with a small number of employees.

4. The Formation of Companies

What sort of a picture does the material give of the formation of enterprises and the associated conditions?

If one follows, year by year, the curve in Fig. 1 for the number of new formations, one can hardly avoid the impression of a surprising stability in the economic situation. The great divergence from this all-pervading relative stability is found during the fantasy years of 1916–20. In these years there was a strong growth in the formation of companies which showed up in the peak years 1918–20 as more than double the formations in the quieter years of 1913–15. The following period of deflation strongly dampened the formation of companies, but it is to be noticed that even during the most difficult years the numbers remained at about twice the level of those before 1916. Thus, in spite of extreme difficulties, the formation of companies appears to continue. The following period of favourable economic circumstances during the latter part of the twenties did not—as one might well have expected—lead to any forcing of new formation. On the contrary, it appears as though there was some relaxation. Thus one would imagine that the number of new formations is determined to a considerable extent by conditions which are not positively correlated with the general economic situation. In the same way the formation of enterprises during the following weak economic period 1931–4 shows a strong power of opposition to the difficult situation. The favourable period which then followed showed no break with the general rule of stability: enterprise

formation continues, but is not increased by the more favourable environment. However, during the year 1940, a year characterized by uncertainty, there was a noticeable movement downwards.

The over-all impression from the 'birth curve' thus seems to be that enterprise formation can be thought to proceed fairly continuously, without being much influenced by the actual economic situation. It is more difficult to draw definite conclusions about the 'strength' of the enterprise formation process since the other element in the comparison, namely the number of 'possible' new formations, is lacking.

5. The Retirement of Companies

The 'mortality conditions'—which are best observed in the percentage curve of Fig. 2—do not show the same stability with regard to the economic situation as do the new formations. During the deflation period at the beginning of the twenties—a deflation already portended in 1918—mortality rises very sharply, and during the 'trough' year of 1921 is sufficient to outweigh the weakened new formations, and thus the total number of companies in active operation, for this one year, shows a decrease. The following favourable years of the twenties produce a reduction in the number of departures, which, towards the end of this decade, falls to the same level as in 1913–16. The 'thinning out'—the term being used without any connotations of causal relationships, but merely as a synonym for 'mortality' and 'retirement'—among enterprises tends thus to weaken with the onset of more favourable economic conditions, but maintains a level of about 4–5 per cent. per year. When the crisis of the thirties occurs the thinning out increases again, but this time does not rise by more than about one-fourth from the best years (1928–9) to the most unfavourable (1932), as compared with an increase of almost two and a half times from 1915 to 1921. The following improvement in the situation presses down on the mortality curve, and during the period 1936–42 it falls continuously to well below 2 per cent.

The over-all impression derived from the 'retirements' curve is of a continuously progressing and, at least during the early part of the period, relatively strong thinning out, the level of which, in the long run, appears to be falling, but with variations owing to sensitivity to the economic situation. One can obtain an indicator for the general strength of the thinning-out process by calculating the time required for half of the companies to disappear. With a mortality rate between 2·6 and 3·6 per cent., characteristic of the period 1935–9, and with the assumptions that new formations and departures are equal, and that the mortality rate applies equally to enterprises in all age-classes, half of the total number of companies at present existing will be renewed in from 20 to 25 years. As a

matter of fact, renewal was carried on at such a rate that the inflow was about twice the outflow, but it was simultaneously counteracted by the fact—which will be shown later—that mortality was higher among younger enterprises than among older.

6. Old and Young Enterprises

One may obtain an impression of the age distribution of enterprise activity between 'old' and 'young' by studying the age distribution of the present enterprises. This distribution is the result of the whole development from a very long time back. In the table below is shown the age distribution actually existing at the end of the year 1942. The groups have been so chosen as to make economic conditions practically the same for the years of formation within each group.

Year of formation	*Number still surviving at the end of 1942*	*Survivors as a percentage of those originally formed*	*Percentage distribution of the total number of existing companies 1942*
Before 1913	2,900	25	14
1913–15	500	28	3
1916–20	1,900	31	9
1921–4	1,700	42	9
1925–30	3,200	56	16
1931–4	2,700	69	14
1935–9	4,200	82	21
1940–2	2,800	97	14
Total	20,000	54	100

The central column in the table shows the total effect of thinning out over the whole lifetime of the existing companies. Less than a third of the earliest group of formations survive right up to 1942. About half of those formed in the twenties have been weeded out, while about 90 per cent. of those from the later thirties and the first war year still remain. If one divides the existing companies in 1942 into three large age-groups, calling those formed before 1916 'old', those from the 15-year period 1916–30 'younger', and the formations still surviving from the thirties and forties 'young', the result shows that almost a fifth are 'old', a good third are 'younger', and something like one-half are 'young'. In general one is fairly safe in assuming that the average size of companies grows with age; thus the share of 'young' companies would have been substantially lower if the measurement had been made in such a way as to indicate the extent of the activity of the enterprises. Probably, though, even after such corrections, the results would not negate the impression that the 'young' enterprises are a quite significant part of private enterprise activity. To a

degree, therefore, the figures appear to confirm the opinion that economic development implies a relatively powerful thinning-out and renewal process going on within the body of existing enterprise.

7. The 'Death-risk' for Limited Liability Companies

It would appear appropriate, by analogy with the census statistics, to interpret the table of the percentage of survivors given above as a sort of 'survivors' table for Swedish companies' giving the probability that a company will reach a certain age. Of those companies formed in the period 1913–15, 28 per cent. remained at the end of 1942. These companies were then between 27 and 30 years old, and the probability of reaching this age could thus be taken to be the figure of 28 per cent.

Such a measure would be correct only if the mortality conditions during the whole period were the same. Actually, however, the intensity of retirement—see Fig. 2—has been different in different years, and in the long run has rather strongly declined. The percentage of survival has thus tended, for the older enterprises, to be lower than it would have been if the average 'mortality rate' for the period had been used, since these 'old' enterprises have been subject to high mortality conditions in the earlier years. In a corresponding way the percentage of survival for the young enterprises is higher than that which would have been obtained if the average mortality conditions for the period had been used, since they have been subject only to the lower 'risk of death' ruling at the end of the period.

If one solves this problem in the way which is usual for tables of survivors in population statistics, one obtains the unbroken line of Fig. 4 for the curve of survivors. This curve is based on the average of all the observations made during the period on enterprises in each age-class.

The diagram gives an impression of the way in which the *average* risk of death ruling during the whole period 1913–42 for each age-class would have successively thinned out the existing group of companies. From 100 new formations during a period, about one-third on the average would have disappeared after 6 years. About 50 per cent. would have been thinned out after another 5 years, and after 22 years from the beginning, not including the first year, about two-thirds would have gone.

One may naturally be hesitant, in a given situation, as to what should be considered 'strong' and what 'weak'. This, of course, has some connexion with what would be 'normal', but it is difficult to develop a 'normal' concept for the retirement of firms. Even if the diagram's curve, owing to the necessarily legal definition of 'retirement', exaggerates the actual thinning out of enterprises, it is most probable that the majority of those who inspect the curve would incline to the view that the thinning out

was 'strong'. The curve may thus, with some justification, perhaps be considered to give some real information about the order of risks to which enterprises during this period have as a whole been subject.

Nevertheless, this survival curve suffers from the weakness that it gives an expression for an 'average' of conditions during years with strongly varying mortality rates. As a measure of the risk of death, it gives too low

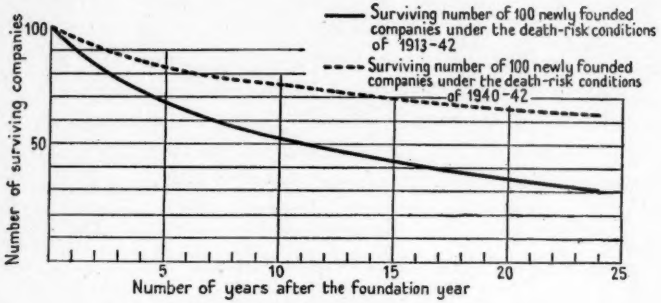
Surviving number of 100 newly founded companies under the death-risk conditions of 1913-42
Surviving number of 100 newly founded companies under the death-risk conditions of 1940-42
Number of surviving companies
100
50
5
10
15
20
25
Number of years after the foundation year

FIG. 4. Companies surviving after 1–24 years.

a picture for the earlier part of the period studied, and for the more favourable years towards the end of the period it strongly over-estimates the power of the thinning out. If, instead, one forms a survival curve solely on the basis of the mortality during the last 3 years, when the mortality rate lay lower than on any other occasion during the period, one obtains the broken curve shown in Fig. 4. This curve clearly differs from the other, and shows that a thinning out of one-third of the original formations now takes 19 years. The conditions during these 3 years also differ markedly from the average for the period.

8. The Extreme Mortality of Young Enterprises

If one studies the survival of death age-year by age-year, it shows clearly that the risk is particularly great during the enterprises' first year, and thereafter successively falls. This is shown, for example, in the survival curve for 1913–42, where the number of deaths in the second year of survival reduce the number of companies in the proportion 91·2 to 83·7, i.e. a good 8 per cent. yearly for companies of this age. The corresponding figures for the fifth survival year are from 44·4 to 42·8, or between 3 and 4 per cent. yearly. Thinning out is thus more than twice as strong in the former age-group than in the latter.

One would have expected *a priori* that the death-risk for companies—as opposed to the death-risk for people—would diminish with age. That a company succeeds in surviving from one year to the next is naturally not, in itself, decisive evidence of increased vitality and strengthened

ability to withstand the economic strain. But in general, and at least up to a certain limit, increase in age is usually associated with financial consolidation, improved position on the market, growing business experience and increased size, eventually with the associated advantages of large-scale operation—all conditions which would be expected to lower the risk

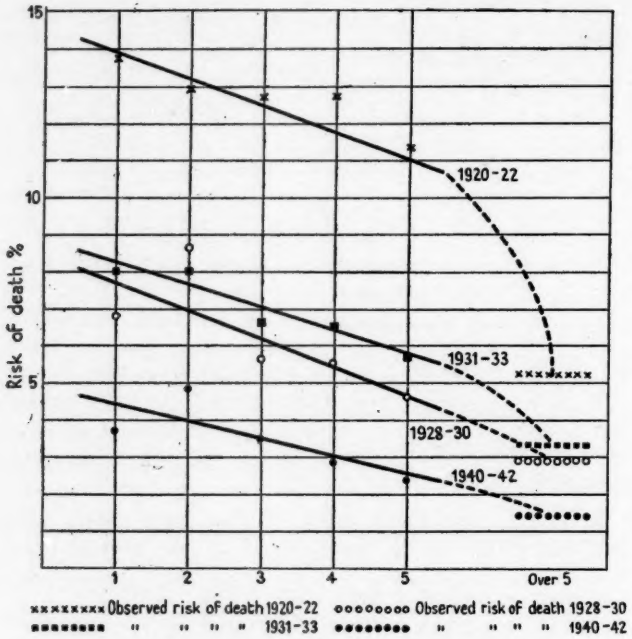

FIG. 5. Yearly death-rate of companies of different ages.

of death. It is clear that only a very few older companies, strengthened by such conditions, achieve the occasionally super-human life-span of the juridical 'person'. As a matter of fact, for the great mass of companies the actual situation appears to be otherwise.

The fact that thinning out is more severe on younger enterprises appears to characterize the whole period. This is illustrated by the curves in Fig. 5.

These curves show the risk of death separately for the first 5 years of survival and as an average for ages above 5 years. Each curve is based on observations of mortality for a special 3-year period. (The points in the diagram show the actual observations to which the straight lines shown are fitted.) The periods are chosen using the percentage of deaths curve of

Fig. 2, and are characterized by mortalities during these periods which are especially high or low in relation to the surrounding years. The four curves thus lie on rather different levels. The mass slaughter of enterprises during the deflation after the First World War shows up strongly in the curve for the period 1920–2 (average mortality 8·3 per cent.). The favourable period of 1928–30 shows a completely different and lower level (average mortality 4·0 per cent.). The moderate increase in the risk of death during the following unfavourable economic conditions shows in the curve for the period 1931–3 (average mortality 4·4 per cent.). The favourable conditions at the end of the period studied show up in the low level of the curve for the period 1941–2 (average mortality 1·9 per cent.). The mortality level is thus rather different for the four periods, but for each such level the mortality of the *youngest* age-class is about double that of the average for the old enterprises. Thinning out is thus mainly concerned with young enterprises.

9. Summary

Strong objections could be raised against the use of this material as a measure of the real turnover in the sphere of limited liability companies. A careful interpretation of the material appears, however, to provide the following general conclusions concerning enterprise formation and death during the inter-war period:

1. New formation of enterprises occurs continually, even during unfavourable economic conditions. The improving economic conditions of the twenties and thirties show no signs of any noteworthy outcrop of new formations.
2. Thinning out of enterprises is relatively powerful all the time, and goes on even during favourable general economic conditions.
3. The rise in the intensity of thinning out was strong during the deflation crisis of the twenties, but was surprisingly weak during the depression of the thirties.
4. The strength of the thinning-out process appears in the long run to have been falling.
5. Thinning out is considerably stronger among the younger than among the older enterprises.
6. The younger enterprises form a relatively significant part of the total number of existing companies.

Since the existing number of companies is dominated by smaller enterprises, the conclusions obtained are mainly applicable to smaller enterprise activity.

UNIVERSITY OF UPPSALA

MARGINAL PREFERENCE AND THE THEORY OF WELFARE[1]

By JEROME ROTHENBERG

I. Introduction

With the contributions of Hicks and Allen in the 1930's the indifference-curve approach to consumer demand came into general analytic usage. Under this approach ordinal properties of utility indices were deemed necessary and sufficient to rationalize all forms of consumer behaviour then incorporated in economic theory. This supplanted the approach stemming from Jevons and Menger, under which analytic conclusions were based on cardinal properties of utility. The empirical implications of cardinal utility were deemed by many economists too controversial to stand at the base of demand theory. If less restrictive assumptions could 'explain' all that was desired of a theory of consumer behaviour, then such assumptions were preferable (by Occam's razor). A sizeable literature was generated by the controversy over whether (1) the indifference-curve approach did indeed eschew cardinal properties of utility, (2) it could 'explain' all that was desired.[2] The outcome of this controversy seems to have been that the indifference-curve (or preference) approach did not involve cardinality[3] and that the approach could 'explain' the empirical content of the then-contemporary demand theory.

Two developments have recently reopened—if indeed it was ever closed—interest in the second issue. First, the von Neumann-Morgenstern *Theory of Games* derives a cardinal utility index by emphasizing the meaningfulness of certain consumer behaviour in the face of risk.[4]

Second (and more relevant to my purpose here), the 'New Welfare Economics', formulated to derive *objective* (non-value saturated) conclu-

[1] I wish to express my gratitude to Professors Albert G. Hart and Ernest Nagel of Columbia University, and Professors James R. Nelson, Theodore Koester, and Mr. David M. Good of Amherst College for their useful comments and criticisms. The author, of course, bears full responsibility for the opinions expressed in this article.

[2] For example, R. G. D. Allen, 'The Determinateness of the Utility Function', *Review of Economic Studies*, 1935; J. M. Clark, 'Realism and Relevance in the Theory of Demand', *Journal of Political Economy*, 1946; Frank Knight, 'Realism and Relevance in the Theory of Demand', *Journal of Political Economy*, 1944; Oskar Lange, 'On the Determinateness of the Utility Function', *Review of Economic Studies*, 1934—to mention but a few of the journal articles.

[3] As employed in this paper, a cardinal index of utility is one for which at least a simple ordering of both preferences *and* preference intensities is defined and intensities are relatable to a common unit. See section II below for a fuller discussion.

[4] But see a proof that 'reasonable' behaviour under risk does not *per se* guarantee cardinality. Only an arbitrary and by no means empirically indubitable kind of 'reasonable' behaviour under risk leads to the von Neumann–Morgenstern index. Baumol, 'The von Neumann–Morgenstern Utility Index: an Ordinalist View', *J.P.E.*, Feb. 1951.

sions on individual and community well-being from ordinal preference theory, has been revealed to possess serious shortcomings, shortcomings which currently prevent it from giving even approximately definitive counsel on any subject.[1] Some of the most important of these shortcomings depend on the lack of cardinal properties of the utility indices assumed for consumers. Cardinal utility indices would significantly strengthen the power and daring of welfare economics.

But does cardinality meaningfully exist? That is, are its implications consistent with observable consumer behaviour; or (not quite identically) can cardinal utility indices be inferred from observable consumer behaviour?

A recent article by W. E. Armstrong, 'Utility and the Theory of Welfare',[2] gives an affirmative answer to this question. His conclusion and his introduction of rather novel ideas in economic analysis merit close scrutiny. My paper will be chiefly concerned with his concept of 'marginal preference' and its implications for welfare analysis.

II

1. Armstrong's important contribution concerns his use of the concept of 'marginal preference' to construct a powerful foundation for welfare economics. Most briefly, 'marginal preference' arises out of an alleged intransitivity of indifference. While preference theory has usually involved (either explicitly or by implication) transitivity of both preference and indifference, a full treatment of the case where indifference is intransitive (but preference transitive) is given by Georgescu-Roegen in his 1936 article.[3] Armstrong's argument in support of intransitivity somewhat resembles Georgescu-Roegen's, but the implications which he draws are essentially different. The following is essentially his argument.

In preference theory, an individual is deemed capable of comparing any three commodity bundles or 'situations' A, B, C, so as to determine between any two whether A is preferred to B, B preferred to A, or A indifferent to B. Moreover, his comparisons are deemed empirically constrained such that: (1) if A is preferred to B, and B preferred to C, then A is preferred to C; (2) if A is indifferent to B, and B is indifferent to C, then A is indifferent to C. (The 'revealed preference' approach to demand theory initiated by Paul Samuelson has been shown in recent articles by Little, Houthakker, and Samuelson to be logically identical with prefer-

[1] See e.g. Kenneth Arrow, *Social Choice and Individual Values*, Wiley, 1950; I. M. D. Little, *A Critique of Welfare Economics*, Oxford, 1950; Paul Samuelson, 'The Evaluation of Real National Income', *Oxford Economic Papers*, 1949, No. 1, and *Foundations of Economic Analysis*, chap. viii.

[2] *Oxford Economic Papers*, Oct. 1951.

[3] 'The Pure Theory of Choice', *Quarterly Journal of Economics*, vol. l, Aug. 1936, pp. 545–93.

ence theory of the Slutsky–Hicks–Allen type,[1] and thus normally to involve transitivity of both preference and indifference. I shall discuss the exceptional instances below.)

On empirical grounds, Armstrong rejects constraint No. 2. There are many instances, he claims, when an individual reports that he is indifferent between A and B *and* between B and C, but *not* between A and C. The behavioural basis of his contention can perhaps be made clearer by extending this example: the individual reports that he is indifferent between A and B, B and C, C and D, D and E, E and F, but not between A and F. In general this means that any relation of preference between A and F can be broken down into a chain of successive indifferences (assuming throughout that the number of such alternatives as the individual can operationally compare—i.e. 'real' alternatives—is significantly large throughout the entire range of choice). Moreover, if we neglect this empirical fact of intransitivity and insist on defining indifference as transitive in our system, we may get outright contradiction. Assume the individual prefers A to A'. Now introduce situations B, C, D, E, and F such that the individual is indifferent between A' and B, B and C, C and D, D and E, E and F. By direct comparison, he would prefer F to A' and F to A. With no such direct comparison, however, we would *infer* through transitivity of indifference that A' was indifferent with F and therefore that A was preferred to F.

Armstrong postulates that intransitivity arises because of the imperfection of human perception. In our original example, comparing A, B, and C, A and B were declared to be indifferent, but they were, in fact—argues Armstrong—felt to be slightly different. The perception of difference involved here was not clear enough to warrant the individual preferring one to the other. C, however, indifferent to B yet also slightly different, was dissimilar enough from A so that a real preference could be perceived. Small, unperceived 'preferences' cumulate into larger, perceivable preferences. In general, then, Armstrong would bring into demand theory considerations about the intensity of preference, a concept not otherwise needed by ordinal preference theorists to deduce consumer behaviour. We shall see below that a systematic treatment of preference intensity is undertaken by Armstrong to derive a 'determinate' ('cardinal') utility function. At this point, however, he uses the concept to redefine the distinction between preference and indifference. The intensity of preference—as an empirical fact—varies over a wide range. No assumption is made as to measurability of intensity, but a simple ordering is deemed

[1] Paul A. Samuelson, 'Consumption Theory in Terms of Revealed Preference', *Economica*, 1948, pp. 243–53; I. M. D. Little, 'A Reformulation of the Theory of Consumer's Behaviour', *Oxford Economic Papers*, 1949, No. 1, pp. 90–99; H. S. Houthakker, 'Revealed Preference and the Utility Function', *Economica*, May 1950, pp. 159–74; but see Appendix.

possible. Some of these intensities are great enough to be called preferences, but some are great enough only to be distinguished from the smallest experienced intensity, not great enough to be considered preference. In other words, there is a threshold in the perception of preference, such that intra-threshold intensities do exist but are not recognized as preference. On this scale of preference-intensities one particular intensity is the greatest of those within the threshold; another is the smallest of those just outside the threshold—the smallest difference perceived as preference. Armstrong calls the first of these 'marginal indifference', the second 'marginal preference'. If preference-intensity is a continuous variable, then marginal indifference is identical with marginal preference. Otherwise they are distinct. Psychological considerations might lead one to question the likelihood of continuity.

The relation of preference and indifference can now be redefined. An individual prefers A to B if and only if he has at least a marginal preference for A over B. An individual is indifferent between A and B if and only if he feels less than a marginal preference between them. And the intransitivity of indifference (but not of preference) can be seen directly. A is indifferent to B, B indifferent to C. But in comparing A with C, the intra-marginal preference between A and B cumulates with that between B and C to give a total preference-intensity which may well be greater than marginal preference. Hence A and C need not be indifferent.

2. How useful is this scheme? Armstrong believes that it lays the foundation for a scientific welfare economics of far greater effectiveness than the 'new welfare economics'. His argument is really twofold. He believes, first, that the notion of intransitivity of indifference alone is sufficient for a more powerful analysis of group-welfare scales than is possible with traditional preference theory. Second, he believes that the concept of 'marginal preference' enables us to formulate a 'determinate' (cardinal) index of group utility under which interpersonal comparisons are possible.

As for the first, how does Armstrong propose to obtain conclusions about group welfare not accessible to traditional ordinal preference theory? In the two-person group, individual No. 1 prefers situation A to B, while individual No. 2 prefers situation B to A.

Now introduce: (1) some situation C which No. 1 orders as follows: A preferred to C, C preferred to B; (2) some situation D which No. 2 orders as follows: A indifferent to D, D indifferent to B, but B preferred to A (because of intransitivity).

(*a*) The situation A is group-preferred to 'a combination' of C for No. 1 and D for No. 2 (since No. 1 prefers A to C and No. 2 is indifferent between A and D).

(*b*) The 'combination' of *C* for No. 1 and *D* for No. 2 is group-preferred to *B* (since No. 1 prefers *C* to *B* and No. 2 is indifferent between *D* and *B*).

Since group-preference is transitive, from (*a*) and (*b*) we infer

(*c*) *A* is group-preferred to *B*.

I believe this demonstration is faulty. The proof depends not only on the validity of indifference in transitivity but also—and crucially—on the meaning of the concept of a 'combination of *C* for No. 1 and *D* for No. 2'. What does this concept mean? Steps (*a*) and (*b*) imply that the 'situations' *A* and *B* can each be compared with a 'situation' in which No. 1 experiences 'situation' *C* and No. 2 experiences 'situation' *D*. Is it possible for such a composite 'situation' to exist? I think not. For individual No. 1, *A* is preferred to *C* and *C* preferred to *B*, so *C* must be comparable with *A* and *B*. Similarly, *A* is indifferent to *D* and *D* indifferent to *B* for individual No. 2, so *D* must be comparable with *A* and *B*. It would appear to be possible for No. 1 to experience *C* at the same time as No. 2 was experiencing *D* only if *C* expressed a distribution of real income to No. 1 alone and *D* expressed a distribution of real income to No. 2 alone. If *C* or *D* included commodities going to both individuals then, in the general case, simultaneous occurrence of both *C* and *D* would be inconsistent. Yet *C* and *D* *must* refer to the real incomes of both Nos. 1 and 2, because both are deemed comparable with *A* and *B*; and *A* and *B* find their way into the preference scales of these individuals. So no meaning can be attributed to Armstrong's composite 'situation'.[1]

[1] This must be qualified. My conclusion is valid when there are external relations in consumption. It is not true only in the absence of all external relations. Thus, if No. 1 is affected by the consumption of No. 2, then a particular flow of goods to No. 1 will have different positions on his scale of preferences depending upon the corresponding real income of No. 2. Let *A* and *B* refer to commodity bundles going to both No. 1 and No. 2, *C* refer to a bundle going to No. 1 alone, and *D* refer to a bundle going to No. 2 alone. Then *C* cannot be uniquely ordered with *A* and *B*, since no real income for No. 2 is specified—*C* is not comparable with *A* and *B* for No. 1. Similarly, *D* cannot be uniquely introduced into a preference scale with *A* and *B*—*D* is not comparable with *A* and *B* for No. 2.

Let *A* and *B* now refer to commodity bundles going only to one person, say, No. 1. Also let *C* and *D* refer to bundles going only to No. I. All four alternatives are now comparable, but now the simultaneous occurrence of *C* and *D* is in general inconsistent.

Armstrong's proof is allowable only if *C* and *D* can refer to bundles going solely to No. 1 and No. 2 respectively, and yet be comparable with *A* and *B*. This is possible if there are no external relations in consumption. The preference ordering of each individual depends solely on goods going to himself. Thus, *A* and *B* refer to bundles going to both, but their ordering depends for each only on the goods going to himself. *C* now refers to a bundle to No. 1 alone, *D* to a bundle to No. 2 alone. The position of *C* on No. 1's preference scale is now uniquely fixed regardless of No. 2's income. The position of *D* is fixed uniquely for No. 2 regardless of No. 1's income. So the composite *C*–*D* situation has a meaning comparable to *A* and *B*—i.e. (*C*–*D*) refers to an alternative of certain goods going to No. 1 and No. 2—and is in general consistent.

But the assumption that there are no external relations in consumption cannot be accepted for application in welfare economics. Cf. Duesenberry, *Income, Saving and the Theory of Consumer Behavior*, Harvard University Press, 1949; also W. J. Baumol, *Welfare Economics and the Theory of the State*, Harvard University Press, 1952.

(It is interesting that Armstrong *can* infer that A is group-preferred to B by his method employing the concept of marginal preference. This will be discussed immediately below.)

In the second part of his argument Armstrong formulates a cardinal index of utility by introducing his assumption No. 5: 'Any individual welfare variation, sufficient to give rise to preference, causes a greater . . . variation of group welfare than any individual welfare variation, sufficient only to give rise to indifference.'[1]

From this assumption, which Armstrong believes as reasonable—and therefore as acceptable—as any employed by ordinal preference theorists,[2] and the concept of marginal preference, Armstrong believes that a cardinal group-welfare index can be deduced logically. Marginal preference becomes the unit by which all preference intensities are measured. But if preference-intensities can be measured, then the alternatives of choice themselves can be scaled quantitatively.[3]

This last can easily be shown. Say that a certain individual orders four alternatives as follows: A, B, C, D (A preferred to B, B preferred to C, &c.). Here we have a simple order of alternatives. In addition, we have a *partial* order of preference intervals (intensities)—i.e. we know that A is preferred to D more than A is preferred to B or to C, or than B is preferred to C or to D, &c.; but we do not know whether A is preferred to B more than B preferred to C or C preferred to D, &c. If we had a complete ordering of these intensities, we would know more about the placement of these alternatives on a figurative choice continuum. If, for example, we knew that A is preferred to B more than B is preferred to C, then we would know that the interval $\overline{AB} >$ interval $\overline{BC}$. Similarly, say $\overline{BC} < \overline{CD}$, and $\overline{AB} > \overline{CD}$.

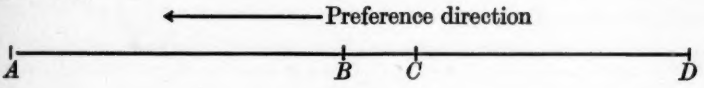

FIG. 1.

Fig. 1 represents one out of an infinite number of placements A, B, C, and D so that $\overline{AB} > \overline{CD} > \overline{BC}$. Numbers could be given to the lateral positions of A, B, C, D, but no arithmetic operations could be performed on such numbers.

If now we had a common unit by which we could compare these intervals, then numbers could be applied to the positions A, B, C, D upon which arithmetic operations could be performed. Thus, taking $\overline{AB}$ as our

[1] Armstrong, p. 264.

[2] One such assumption, as quoted early in Armstrong's article, is: 'If an individual prefers alternative A to alternative B, then his individual welfare is higher with A than with B' (p. 259). In other words, this is the assumption of consumer sovereignty.

[3] See Appendix applying marginal preference to ordinal preference theory.

unit, if it were operationally possible to determine that, say, $\overline{AB}$ = three times as large as $\overline{BC}$ and two times as large as $\overline{CD}$, then the interval $\overline{AB}$ could be given the number 1, $\overline{BC}$ the number $\frac{1}{3}$, $\overline{CD}$ the number $\frac{1}{2}$. And A, B, C, D could be given the numbers 0, 1, $1\frac{1}{3}$, $1\frac{5}{6}$, or 4, 5, $5\frac{1}{3}$, $5\frac{5}{6}$, &c , or any linear transformation of these. These numbers could be added or subtracted, since the intervals remain invariant under such operations. (If, in addition, we knew that particular alternative X for which we could, with reason, give the number zero, then the resulting scale—a ratio scale —would involve measurement on which all the operations of arithmetic could be performed.)[1]

Applying this to marginal preference as the common unit, we can construct an interval scale of utility (where intervals between alternatives can be reduced to multiples of the common unit), and can quantitatively measure the intensity of any preference.

For the first, take some alternative A. Now choose some B, C, D, E, &c., such that B is marginally preferred to A; C is marginally preferred to B; D marginally preferred to C; &c. So $\overline{AE} = 4 \times \overline{AB}$ and if $\overline{AB}$, the marginal preference interval, is given a number, say 1, then $\overline{AE} = 4 \times 1 = 4$ and E has the position $(4+A)$. Again, it must be pointed out that the zero point and the number given to the common unit are arbitrary. For the second, we are given that some E is preferred to some A, and we desire to quantify the intensity of the preference. Now, choose some $B, C, D, \ldots, N$ for which B is marginally preferred to A, C marginally preferred to $B, \ldots (N-1)$ marginally preferred to N, and N indifferent with E. The intensity $\overline{AE} = N \times AB = N \times$ marginal preference.

3. According to Armstrong, marginal preference as such a unit of measurement for individual and group welfare indices, follows logically from assumption 5.

> Since [he asserts] between two alternatives, A and B, A is preferred to B or B is preferred to A or A is indifferent with B, must hold for an individual, and since group-welfare rises with increase of individual welfare, there must be a minimum individual welfare-difference consistent with preference, for each and every member of the group, and this minimum individual welfare-difference must, for every individual, give rise to the same group-welfare relation. If we use the same number, e.g. unity, as a measure for this minimum for all individuals, a group-welfare function emerges having the simple characteristic that the same group-welfare variation results from any given individual welfare variation. In other words, group-welfare variation is a rising function of the sum of individual welfare variations. Let us call the above preference a 'marginal preference'.[2]

And to make explicit how interpersonal comparisons of utility are involved in his index, he argues that

> a marginal preference indicates the same welfare-difference, independently of the person exercising the preference. . . . A marginal preference of any person has a

[1] Armstrong's article even proposes a solution to the problem of setting a zero point.

[2] Ibid., p. 265.

greater effect on group-welfare than a marginal indifference (in the same direction) of the same or any other person by assumption 5. But a marginal preference is adjacent, for any given person, to a marginal indifference (in the sense that either is the boundary of the other). If, therefore, a marginal preference of one person has a greater effect on group-welfare than a marginal preference of another person, then the former's marginal indifference cannot have a smaller effect. But this would contradict assumption 5. Therefore, a marginal preference of one person must have the same effect on group-welfare as a marginal preference of any other member of the group.

If we assume that a variation of group-welfare is a function of individual welfare variation, it follows that a marginal preference implies an individual welfare difference that is independent of the individual. *In other words, interpersonal comparisons of utility are logically implied by assumptions that economists do not seem to question.*[1]

This is quite a remarkable conclusion, all the more remarkable if it can, indeed, be deduced from assumption 5 and the definition of marginal preference. Moreover, if this conclusion does hold by logical implication, then effective criticism must be largely criticism of the seemingly innocuous assumption 5. *Does* cardinal utility follow deductively?

Assume a two-person group and an initial situation A. Now, from A generate a situation B which individual No. 2 marginally prefers to A, but to which individual No. 1 has an intra-marginal preference (i.e. is indifferent between A and B, which can, of course, mean that A is intra-marginally preferred to B). Assumption 5 says that B is socially preferred (i.e. has a higher group welfare) to A. But this conclusion follows also on strictly ordinal preference grounds, since no one *prefers* A to B, and at least one person *prefers* B to A.[2] Similarly, from A generate C which No. 1 marginally prefers to A, while No. 2 is indifferent between A and C. By assumption 5, C is socially preferred to A; and this shows that a marginal preference by individual No. 1 is 'as good as' a marginal preference by individual No. 2 in generating *social* preference. But again this conclusion follows from traditional ordinal preference theory. Does assumption 5, then, take us any farther than traditional theory? Does it, specifically, take us as far as Armstrong asserts?

Let us extend our example. From B generate D such that No. 2 marginally prefers D to B and No. 1 is indifferent between D and B and also between D and A. From D generate E such that No. 2 marginally prefers E to D and No. 1 is indifferent between E, D, B, and A. Now compare

[1] Armstrong, p. 266, footnote 1. Italics mine.

[2] The rules are that:
(1) where no one prefers A to B but at least one person prefers B to A, then B is socially preferred to A;
(2) where everyone is indifferent between A and B, A and B are socially indifferent;
(3) where at least one person prefers A to B and at least one prefers B to A, A and B are socially indifferent. (Majority rule, weighted voting, are exceptions.)
Rule 3 is sometimes rejected, any A and B fulfilling its conditions being considered not comparable. In this case, group-welfare comparisons would enable only a partial ordering of alternatives.

E with C. On ordinal preference grounds alone, No. 1 is indifferent between E and A. He prefers C to A. Allowing for intransitivity of indifference, either C preferred to E or C indifferent with E.[1] Individual No. 2 prefers E to D, D to B, B to A, so E to A (preference is transitive). He is indifferent between A and C. So either E preferred to C or E indifferent with C. As a result of these individual orderings, either C is socially preferred to E (No. 1 prefers C to E, No. 2 indifferent), E is socially preferred to C (No. 2 prefers E to C, No. 1 indifferent), or they are socially indifferent (both indifferent, or opposite choices).[2]

Now apply Armstrong's cardinal utility index to this comparison. The sum of the preference intensities of C to A is:

$\overline{CA}_1 = 1$ (= marginal preference, arbitrarily numbered 1) preference intensity of individual No. 1

$-1 < \overline{CA}_2 < 1$ (= non-transitive indifference between C and A) preference intensity of individual No. 2

$\overline{CA}_1 + \overline{CA}_2 = \overline{CA}_{1+2}$, where $0 < \overline{CA}_{1+2} < 2$.

The sum of the preference intensities of E to A is:

$-1 < \overline{EA}_1 < 1$ (indifference)

$\overline{EA}_2 = \overline{ED}_2 + \overline{DB}_2 + \overline{BA}_2 = 1+1+1 = 3$ (sum of three marginal preferences)

$\overline{EA}_1 + \overline{EA}_2 = \overline{EA}_{1+2}$, where $2 < \overline{EA}_{1+2} < 4$.

Thus $\overline{EA}_{1+2} > CA_{1+2}$, so E is socially preferred to C.

Finally, what do assumption 5 together with the concept of marginal preference tell us? They tell us that individual No. 2 cannot be indifferent between C and E; only a preference for E over C is consistent with his ordering. A and C are indifferent, but B is preferred to A; so by assumption 5, B involves a greater individual welfare-variation than C. Then, since D is marginally preferred to B, and E to D, we know that the preference intensity of E to C is greater than marginal preference, so E is preferred to C. This individual preference then tells us that C cannot be socially preferred to E.

[1] E preferred to C is precluded. If E preferred to C, then, since C preferred to A, E preferred to A. But E indifferent with A. So not E preferred to C.

[2] Note that we cannot employ the following analysis:

(1) comparing B to C, No. 1 prefers C to B, No. 2 prefers B to C, so B and C are socially indifferent;

(2) comparing E to B, No. 2 prefers E to B, No. 1 is indifferent between them; so E is socially preferred to B; and thus

(3) E is preferred to C.

With intransitivity of indifference, (3) does not necessarily follow (1) and (2). And in fact we know that a direct comparison between E and C shows that No. 1 prefers C to E, No. 2 prefers E to C, so they are socially indifferent.

This inconsistency illustrates Arrow's doubts about acceptable social welfare functions with a vengeance!

Can they tell us more? Introduce L such that No. 1 is indifferent between L and C and also between L and A (thus A indifferent to L, L indifferent to C, and C preferred to A); and No. 2 marginally prefers L to C. So, L is socially preferred to C. Now compare L with B. Since No. 1 is indifferent between A and B, he must either prefer L to B or be indifferent between them.[1] Individual No. 2 marginally preferred B to A and was indifferent between C and A, so No. 2 is indifferent between B and L.

Continue by comparing E with L. For No. 1, E is indifferent with B and A; consequently, either L preferred to E or L indifferent with E. For No. 2, E is marginally preferred to D which, in turn, is marginally preferred to B. Thus, by assumption 5, D represents a larger welfare variation from B than does L, and since E marginally preferred to D, the preference intensity of E over L is greater than marginal preference.

If No. 1 were indifferent between E and L, then, E would be socially preferred to L, L socially preferred to C, so E socially preferred to C. Using assumption 5 and the concept of marginal preference would give more information than ordinal preference and as much as cardinal utility. But No. 1 is not in general indifferent between E and L. We cannot throw out the case where No. 1 prefers L to E. And with this case, we are back where we started in the comparison: No. 1 prefers L to E, No. 2 prefers E to L.

A further point must be made. Throughout this final extension of our example we have proceeded as though marginal preference and intramarginal preferences were invariant distances. We have inadvertently attributed cardinal properties to these concepts. It was these properties that enabled us to derive more information out of assumption 5 than out of strict preference theory. Thus, in our above analysis: 'A and C are indifferent, but B is preferred to A; so by assumption 5, B involves a greater individual welfare variation than C. Then, since D is marginally preferred to B, and E to D, we know that the preference intensity of E to C is greater than marginal preference, so E is preferred to C.' Here we are adding numerically invariant magnitudes, a process which we cannot infer from assumption 5.

All assumption 5 does, in fact, is to suggest rules by which a cardinal unit of measurement, *already introduced into preference*, may be specifically manipulated. It is not this assumption that leads to cardinality. It is rather the assumption that marginal preference is an appropriate description of preference behaviour. *Without marginal preference, assumption 5 tells us nothing which strict preference theory does not tell us. Without*

[1] He cannot prefer B to L, since considering how L was chosen, that would imply that he preferred B to A, which he does not.

assumption 5, marginal preference would none the less give rise to a cardinal index of utility.

All this concerns the appropriateness of marginal preference in the behaviour of a single individual. The reader will notice that even assuming such appropriateness in my example above did not suffice to secure as much information as was possible by Armstrong's numerical index. What was lacking was an additional assumption about the comparability of the marginal preferences of different individuals. Assumption 5 does not guarantee such comparability. Marginal preferences are not called on to effect group-welfare directly, but rather *through the welfare of individuals.* Only an additional explicit assumption as to interpersonal comparability of welfare changes can lead to Armstrong's numerical index.

To conclude this section, a cardinal utility index allowing for interpersonal comparisons of utility does not follow logically from assumption 5. It does follow from these additional (implicit) assumptions:

(1) marginal preference is an empirically observable, relatively invariant, element in the choice behaviour of individuals;

(2) the marginal preferences of different individuals are directly comparable in relatively invariant fashion.

III

That a cardinal utility index cannot be deduced from assumption 5 does not mean that such an index is therefore useless. On the contrary, if the implicit assumptions made explicit at the end of the previous section should be warranted (if, truly, they are the assumptions that are required), then a significant contribution will have been made. A firmer, broader foundation will indeed have been constructed for welfare economics. But it appears to the writer that, considering the purposes for which they are evoked, neither of these implicit assumptions *is* warranted.

1. The first assumption really has two parts: (*a*) that marginal preference (implying the intransitivity of indifference)[1] is empirically observable; (*b*) that it is relatively invariant. Throughout my paper I have proceeded by tentatively assuming that (*a*) is valid. Clearly the matter cannot be so decided. Far more direct empirical work must be done to determine whether there *is* significant intransitivity of indifference; under what conditions it appears or fails to appear; whether marginal preference *is* an important facet of human choice; whether it varies, and why, and how much.[2] Logical constructs built up largely out of introspection

[1] But not conversely. See my argument below.

[2] For interesting data on this, obtained by empirical derivation of a composite preference scale with two commodities, see Stephen W. Rousseas and Albert G. Hart, 'Experi-

cannot be depended on as foundations for a theory with such far-reaching implications as this one. The remarks that follow continue to assume that marginal preference exists, and apply what less-than-definitive information there now is available to throw light on what properties we might expect it to possess.

2. A discussion of preference intensities, especially of intra-threshold intensities, must recognize the dependence of preference on the nature of perception. Marginal preference depends crucially on perceptions of sameness and difference, comparisons not simply between quantities of a single commodity, but between complexes of many different commodities. As noted above, in order to be an appropriate unit of measurement, marginal preference must either be nearly invariant or must vary in an orderly (highly predictable) fashion. But the perceptions on which marginal preference depend are far from uniform for different people, or even for the same person. The level of discernment depends on individual constitution, motivation, the degree of learning of the individual, and even on the momentary state of the organism (for example, fatigue, elation, &c., will affect such perceptions). For a given individual, his level of perception will differ for complexes with different kinds of items,[1] since his motivation towards them and the degree of his knowledge about them will differ. Thus, to take an extreme example, if an individual marginally prefers B to A and marginally prefers C to A, then he may nevertheless prefer B to C, because the A–B and A–C comparisons are psychologically independent and do not determine a still-different comparison between B and C. Marginal preference becomes a different thing for different comparisons, even comparisons taken at the same period of time.[2]

Moreover, an implication of this relationship between marginal preference and perception is that marginal preference can be expected to change frequently over time, more frequently than preference orderings, since a given ordinal preference scale is compatible with many different marginal preferences, but very likely not conversely. Can Armstrong's theory of marginal preference effectively deal with changes over time?

One suggestion might be to treat future perceptions of preference the

mental Verification of a Composite Indifference Map', *Journal of Political Economy*, Aug. 1951, pp. 288–318.

[1] The fiction that every consumer consumes some of every commodity available must be dropped to achieve a degree of realism consistent with the discussion of marginal preference.

[2] Notice that while my example is not consistent with the concept of a unique marginal preference, it is consistent with traditional preference ordering. It may, of course, turn out that empirical studies emphasizing the diversity of choice will find even significant *ordering* inconsistencies, so that in some instances only partial orderings of preference can be accepted. As mentioned in my Appendix, much of the problem about the specification of preference scales under intransitivity of indifference depends on whether finitude of levels of discretion (and welfare) is a more important phenomenon than fuzziness in different comparisons with a particular alternative.

way future goods are treated in Hicksian comparative statics. On inspection, however, this scheme appears inapplicable. Present and future commodities can meaningfully be held within a single preference scale (e.g. there is a real sense in which present and future satisfactions are substitutable). But what can be the meaning of simultaneously contemplating a present perception of some two combinations (say A and B) and a future perception of the same two combinations? If the future perception can be conceived of in detail in the present, then it is part of the present perception. Changes over time in neither preference orderings (tastes) nor in perceptions of marginal preference can be included within the preference frame of reference of a single point of time.

The situation is even more serious for changes in marginal preference than for changes in tastes. The latter simply involve reordering alternatives, and the new and old preference scales are comparable. But changes in marginal preference mean changes in the whole frame of reference of utility scales. The new scales are by no means obviously comparable with the old. So consistent temporal comparisons of utility change over time are most difficult.

The important consequence of these changes over time is linked to my doubt concerning the reality of a unique measure of smallest perceivable preference between vastly dissimilar multi-dimensional complexes at any *one* time. For changes in marginal preference would not be necessarily uniform for different pairs of alternatives. At any one time the smallest perceivable preference between any one pair of alternatives may be *defined* as identical with that between any other pair. Then, if changes over time affect the comparison of every pair similarly, we may still be permitted to speak of 'a' marginal preference—albeit a changing one—since the original internal structure of preference comparisons will be unchanged. But if changes over time disrupt the original structure, and in largely unknown ways, then it is indeed questionable whether we can talk meaningfully about a concept which is assumed to have measurable characteristics yet at no time can be uniquely specified.

It might be objected that variations in marginal preference are not necessarily fatal so long as these variations can be functionally explained. True; but it is not sufficient simply that changes in the unit of measurement be 'explainable' by a less proximate set of variables. If these variables are themselves unpredictable, and especially if they give rise to frequent irregular fluctuations of—and within—marginal preference, nothing is gained by the explanation and we have a unit of measurement which is no unit at all. A more useful explanation—necessarily quantitative—does not exist with the current level of psychological understanding.

The above comments refer only to marginal preference for a single

individual. If marginal preference is so crude in this relatively simple context, what likelihood that we can confidently compare marginal preferences for *different* individuals ?[1]

In conclusion:

1. Acceptance of non-transitivity of indifference neither logically nor empirically requires acceptance of the concept of marginal preference. We can have the first without the second, although not conversely.
2. Armstrong has not succeeded in proving—to my mind, at any rate—that ordinary preference analysis, supplemented only by non-transitive indifference and assumption 5, itself implies a group utility function with cardinal properties.
3. The concept of marginal preference as so far developed does not appear to lend itself to a useful formulation of cardinal utility, either for individuals or for groups. It does certainly not at this time provide the key to interpersonal comparisions of utility.

APPENDIX

The Application of Marginal Preference to Ordinal Preference Theory

Modifying traditional preference theory with Armstrong's considerations is not easy. Two general courses suggest themselves. We can translate marginal preference and the various degrees of intra-threshold preference into a band of fuzziness, or penumbra (1) around each indifference curve; or (2) around each alternative of choice. With either course, an implication is that an individual's levels of welfare discrimination are decreased sharply to some finite number (in the realistic case of a saturation area; cf. J. M. Clark, op. cit., pp. 349–50). But these reductions lead to difficulties.

Under the first course, it might be expected that a possible interpretation is one in which the indifference curves are viewed as step-functions, the step interval being equal to an assumed constant marginal preference. This interpretation, however, leads to a contradiction, since it will not necessarily be true that a combination at one welfare level is at least marginally preferred to any combination in the next lowest welfare level. A preference intensity less than marginal preference can give rise to a welfare change of marginal preference. To see this, consider the diagram on p. 262.

For graphical purposes we may simplify our alternatives of choice into two-dimensional vectors—i.e. combinations of commodities X and Y. Our curves do not here specify particular preference levels, but rather each pair of curves expresses the boundaries of all combinations of A and B which are perceived as giving the same satisfaction. All combinations within the area I are equally satisfactory.

[1] Such comparisons would require not only that marginal preference be meaningful for one individual, but also that 'the' marginal preference of individual A stood in some relatively constant relationship to 'the' marginal preference of individual B over time, so that they could at least be *defined* equal without such definition being uselessly arbitrary. Our consideration of marginal preference for one individual suggests strongly that : (1) we would have little knowledge about the original relationship, and (2) we would not be deluded into feeling that the relationship was very likely to remain undisturbed.

Similarly for II, each combination of which is deemed marginally preferred to any combination in I.

Let us take distances in the preference direction (i.e. lengths of the gradient vectors) as representing preference intensities. For this, we must assume a continuum of continuous indifference curves within I and II. The difference between any two such indifference curves within the same perception band (as given content

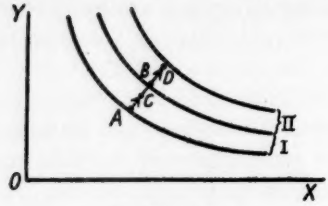

DIAGRAM 1.

by the non-transitivity of indifference which we are temporarily assuming throughout section II) represents intra-marginal preference intensity. Measuring such distances is, of course, tolerably proximate only if we confine ourselves to sufficiently small areas. But since our discussion need concern only a few adjoining perception levels (out of a much larger totality), this condition is met both for intra- and extra-marginal intensities involved.

An alternative *B*, just within the lower boundary of II, must be marginally preferred to alternative *A*, just within the lower boundary of I (and with *B* on *A*'s gradient vector, for simplicity). So marginal preference must be a preference intensity at least as great but no greater than preference distance $\overline{AB}$: if it were greater, *A* and *B* would be indifferent; if it were less, some *C* within I would be marginally preferred to *A*, which is a contradiction.

Let preference distance $\overline{CD}$ equal $\overline{AB}$ (where *C* is in I, *D* in II, and both are on *A*'s gradient vector). *D* is marginally preferred to *C*. But *B* is also marginally preferred to *C*, because they are in adjoining perceptible preference levels, by definition. And $\overline{CB}$ is less than $\overline{CD}$. Hence, in this interpretation, a preference intensity less than marginal preference can give rise to a welfare change of marginal preference, if marginal preference implies the existence of intra-marginal preferences.

A variant of this, whereby we assume that the upper boundary of I is separated from the lower boundary of II by a preference distance equal to marginal preference, might avoid this contradiction. But it could not explain how *B* could be *just* marginally preferred to *A*, unless no preference intensities *within* I were admitted. In other words, step-function indifference curves can express constant preference distances only when intra-marginal preferences do not exist—when indifference is transitive. But with transitivity of indifference the preference distance between indifference levels can no longer be interpreted as *marginal preference*.

If we wish to retain the concept of marginal preference, the above dilemma can presumably be reconciled either if marginal preference is not held to be a constant preference distance (which the realism of finite fuzzy levels of discrimination makes additionally attractive; also see the last section of this paper for a more searching critique of its 'constancy'), or we must adopt the second major course of approach mentioned above, treating every *alternative* rather than every welfare level as having its area of intra-threshold preference. A suggested illustration of this is given by the diagram on p. 263. All alternatives within a given circle are indifferent with the focus alternative (the centre of the circle); but indifference is not transitive—thus *A* is indifferent with *C*, *C* is indifferent with *B*, but *A* is not indifferent with *B*. Note, too, that the above contradiction does not apply here. The area of intra-marginal

preference might approximate the more familiar tubular shape convex to the origin only if one could specify preference directions; but this would necessitate each alternative being completely indifferent (absence of even intra-marginal preference) with a set of sufficiently many other alternatives. The more tubular the shape, however, the closer we would be approximating the step-function interpretation,

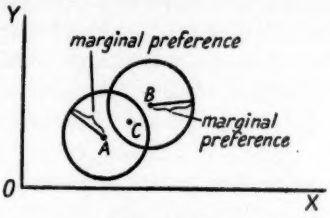

DIAGRAM 2.

and progressively less would be gained by treating each alternative rather than each welfare level as a focus.

This second general interpretation of marginal preference seems closer to Armstrong's emphasis on non-transitivity of indifference (and to the revealed preference approach), but its full implications, other than a possibly significant range of indeterminacy in predicting consumer behaviour, are not clear at this writing.

One last point should be made. With significant non-transitivity of indifference, the Hicks–Allen indifference approach seems translatable into the first but not the second major interpretation type, while Samuelson's revealed preference approach seems translatable into the second but not the first. Only if indifference classes are large do the two interpretations become very similar. Under strong non-transitivity of indifference—with or without constancy of marginal preference—indifference curve and revealed preference analysis do *not* appear to be logically equivalent.

AMHERST COLLEGE

A REPLY

By W. E. ARMSTRONG

HAD Rothenberg's exposition of my theory been less fair and his criticism more antagonistic, the task of commenting on his article would have been less difficult. In so far as he denies that my conclusions follow from my axioms, the matter is straightforward. But, as I hope to show, his main criticisms are really disguised objections to my axiom (5),[1] though he discreetly avoids openly rejecting this axiom. It would seem, therefore, that what I must do is to show that Rothenberg's standpoint *does* involve rejection of axiom (5), and to bring into the open my own standpoint which allows this axiom to survive.

Let me deal first with what I feel is the minor part of his article; the suggestion that my conclusions do not follow from my axioms. I think I may take it that there is a tacit agreement to ignore two amplifications of my theory: the determination of zero individual welfare, and the qualification for equity. The question then is, Are my deductions from axioms (1), (2), (3), and (5), the assumption of transitive preference and the denial of transitive indifference, logically sound?

I have to admit, at the outset, to the omission of an important step in my argument, referred to, with some justification, by Rothenberg (p. 252 of his article) as a faulty demonstration. My axioms do not rule out 'external consumption relations' and I have no wish to do so. There can be no doubt that α's preference for a situation may be affected by his beliefs about β's consumption, and, similarly, β's preference by β's beliefs about α's consumption. If α's and β's beliefs are true, then C and D may be a logically impossible combination (e.g. C may be 'consuming x by α and consuming y by β' and D may be 'consuming x' by α and consuming y' by β', and $x \neq x'$ or/and $y \neq y'$). I omitted to point out that α's belief in β's consumption is logically consistent with any consumption in fact performed by β, since beliefs can be false—similarly for β's belief in α's consumption. Since α's and β's beliefs about each other's consumption need not be true, C and D become logically possible combinations, and the demonstration to which Rothenberg objects remains valid.

Even if an extreme view be taken about external consumption relations (i.e. that α's preferences depend upon the *fact* of β's consumption, rather than on α's *belief*—a view that seems to me clearly false), the difficulty is

[1] My formulation of axiom (5) was as follows: 'Any individual welfare variation, sufficient to give rise to preference, causes a greater (or "wider", if group-welfare is indeterminate) variation of group-welfare than any individual welfare variation, sufficient only to give rise to indifference.'

only apparent. The fact that, on this view, it may be logically impossible to combine the situations C and D does not imply that it is logically impossible to have a combination of 'welfare of α corresponding to C' and 'welfare of β corresponding to D'. On the contrary, any combinations of individual welfares that are separately conceivable are logically possible, though it might be claimed that logical rigour needs an axiom to this effect. If this be granted, then by axiom (5) a group-welfare is conceivable that is lower than the group-welfare of A and higher than the group-welfare of B; and, therefore, the group-welfare of A is higher than the group-welfare of B.

Rothenberg then proceeds to an exposition of my theory (i.e. of what I assert can be deduced from the assumptions enumerated), and I have little fault to find with what he attributes to me. I would, however, take exception to his statement on p. 254 that I deduce from axiom (5) a cardinal group-welfare index. I have been careful to insist that axiom (5) is a weak form of the stronger axiom that is the foundation-stone of the 'new welfare economics', and it is, of course, only an ordinal index of group-welfare that can result from the axioms of that school. I explicitly introduced axiom (6), at a later stage in the argument, in order to pass from an ordinal to a cardinal group-welfare index. However, the point is not important for the exposition that follows, and it is not until p. 257 that I need take issue with Rothenberg. I fully agree with his italicized remark about axiom (5)—of course axiom (5) tells us nothing that preference theory, as usually presented, does not tell us; which is why I can safely claim that school's assent to it as an axiom: and, of course, marginal preference, with or without axiom (5), gives rise to a cardinal index of utility.

In the paragraph that follows we come to the crux of the matter. He asserts that what I require (p. 258) is an 'additional assumption about the comparability of the marginal preferences of different individuals. . . . Only an additional explicit assumption as to interpersonal comparability of welfare changes can lead to Armstrong's numerical index.' And he concludes that two additional assumptions are needed (last two paragraphs of section II), namely:

(1) marginal preference is an empirically observable, relatively invariant, element in the choice behaviour of individuals;

(2) the marginal preferences of different individuals are directly comparable in relatively invariant fashion.

I disagree. We may look at what I have built on my axioms as emerging in two steps.

1. Utility is cardinal for a given individual, and so is his welfare, and there is a one-one correspondence between his utility and his welfare. I do not think that Rothenberg disputes this consequence of my axioms. It

then follows, by axiom (5), that a marginal preference always indicates for that individual the same individual welfare variation (or, to be meticulous, the same group-welfare variation for the group consisting only of that individual), for only if that were the case would an individual welfare variation associated with preference always exceed an individual welfare variation associated with indifference. An individual welfare index thus results, given an arbitrary choice of zero for some welfare position, and an arbitrary choice of a number (e.g. unity) for the welfare variation corresponding to a marginal preference.

It must, however, be admitted that implicit in this conclusion is the assumption of continuous variability of utility and, therefore, of welfare, and that an axiom to this effect might well have been included in my list.

2. By using, for every member of a group, the same index for an individual welfare difference corresponding to a marginal preference, an index of group-welfare variation results by simple summation of individual welfare variations, by axiom (5) alone. This index is ordinal, but becomes cardinal by the introduction of axiom (6). That axiom (5) alone justifies the step taken is easily seen. If, for some combinations of variations of individual welfares, group-welfare did not vary according to the above function, it would follow that somewhere an individual welfare variation had a greater effect on group-welfare than the same individual welfare variation (when measured in the way indicated above) of some other member of the group. But that would mean that an individual welfare variation, associated with an indifference, had a greater effect than some other individual welfare variation, associated with a preference, and therefore axiom (5) would be contradicted.

It is, of course, important to realize that, by the above method, we obtain only an index of group-welfare variation. Each individual welfare index has its own arbitrary zero, and therefore for a group of n persons the zeros chosen determine as many alternative zeros for our group-welfare index as there are ways of combining the choice of n individual zeros. The group-welfare index I proposed was offered explicitly with that limitation. It is one that follows from my axioms and does not require the additional assumptions that Rothenberg deems necessary. Rothenberg's subsequent denial of these additional assumptions is, therefore, irrelevant for the purpose of refuting my thesis.

That I do not need the two additional assumptions cited by Rothenberg does not mean that I wish to deny the truth of these assumptions. On the contrary, the starting-point in my own mind was in fact assumption (2) or something very like it. But when I found that an axiom already basic to the 'new welfare' approach would give me what I needed, I succumbed to the temptation of attacking this school with one of its own weapons.

In section III Rothenberg proceeds to dispute the validity of the above additional assumptions, but, what is more important, he implicitly disputes the validity of axiom (5). It is here, to my mind, that the important criticism of my theory lies. If doubt can be sown about axiom (5), there is little to be said for the theory, even though there may, as a consequence, be even less to be said for the 'new welfare theory', which also requires the truth of axiom (5). I did not feel called upon in my article to exercise any persuasion in favour of this axiom, in view of its apparently universal acceptance; and it is significant that Rothenberg does not openly dispute it. Rothenberg's doubts about what I had assumed to be universally accepted go even farther. He does not seem prepared to accept as universal the transitivity of preference, though here again he does not come out into the open. I had seen no reason for putting up a case for this assumption. After all, both preference and indifference have usually been taken to be transitive; my arguments against such transitiveness for indifference should do nothing to shake the generally accepted view that preference is, without qualification, transitive.

Section III makes clear the standpoint taken by Rothenberg. It also makes clear why he maintains that a number of questions require answers based on empirical investigation before my theory could be acceptable. My standpoint is quite different and when properly understood explains why the questions posed by Rothenberg are for me meaningless. Let me then go into this matter which lies outside the scope of my article and try to justify the standpoint there adopted, i.e. to indicate why I stand by the assumptions that lie behind the theory as well as insisting that the theory follows from the assumptions. As a first step let me bring into the open what is implicit in Rothenberg's standpoint.

At a given moment, so I interpret Rothenberg, an individual can be assumed to be endowed with a utility function which is relatively (but not necessarily completely) independent of his thinking at that moment, so that if he thinks of alternatives A and B he becomes aware of the utility of A and of B. In addition to a utility function that may be changing from moment to moment, an individual is variable in other ways such as in the acuity of his discrimination. He might, therefore, at one moment prefer A to B and at another moment be indifferent between C and D, even though the utility-difference between A and B is less than the utility-difference between C and D. This could happen if his power of discrimination had declined from the earlier to the later moment. Even at one and the same moment, when an individual compares A with B and also C with D, it cannot be ruled out that his discrimination works differently in the two comparisons. Clearly, if we correlate individual welfare with the individual's utility function, the concept of marginal preference

does not help to throw any light on an individual's welfare—not even on the welfare-difference between a pair of alternatives. Only if there were some way of showing that marginal preference is invariant (or how it varies, if we could find some way of measuring an individual's acuity of discrimination at any given moment) would it be possible to construct an index of individual welfare. The most we could ever say of an individual is that his welfare rises from B to A if he prefers A to B.

Furthermore, even if we could obtain a partial solution of this problem for the individual (e.g. by discovering upper and lower limits to the variations of his powers of discrimination), there could be no solution to the problem of comparing group-welfares, if group-welfare is to be a function of individual welfares, since there could be no basis for the assumption that marginal preferences of different persons indicates anything whatever about their relative welfares. Since no empirical investigation could conceivably prove the invariance of marginal preference between persons, the most we could ever say of two situations A and B for a group would be that if every member of the group prefers A to B then group-welfare is higher with A than with B. Finally, it should be noted that axiom (5), far from being 'apparently innocuous', as Rothenberg describes it, is patently false.

Given such a standpoint, I concede all that Rothenberg says and I ask only that he stoutly deny axiom (5) and include in his rejection of my theory an unqualified rejection of the 'new welfare-economics', and that he warn all welfare-economists against any attempt to link the preferences and indifferences of individuals with the welfare of the group in any significant manner.

In rejecting the above standpoint I nevertheless fully agree that in a mental process we must distinguish between the act of discrimination and what is presented for discrimination. I have been at some pains to stress this point in earlier articles,[1] and it is one facet of the argument for cardinal utility. Psychologists have repeatedly emphasized the point in explaining the non-transitiveness of the relation 'not perceptibly different' in judgements about the constituent parts of fields of perception. But that is not to say that a change in the act of discrimination is in any sense a less significant change in an individual's experience than is a change in the objects presented for discrimination. Consider, for example, an individual presented with a pair of sense-data of colour, eliciting the judgement that they are indistinguishable. We know that they differ as sense-data if, when comparison is made with a third colour, this is judged to be indistinguishable from one but distinguishable from the other. Thus there are two ways in which we could change the experience (of no perceptible

[1] e.g. 'Determinateness of the Utility Function', *E.J.* 1939.

difference) to one that elicits the judgement that they are just perceptibly different. One way would be to vary one of the sense-data until the required judgement appears. The other way would be, if only we knew how to do it, to increase the individual's acuity of discrimination until he judged the two colours, that are in fact different as sense-data, to be 'just perceptibly different'. Could it be claimed that the experience thus obtained is the same experience, in spite of the difference in the way it is constituted? I do not know the answer and I have some doubts about the legitimacy of posing the question, but at least the question is not pointless. There is something about an experience of 'just perceptible difference' which has significant identity however it is constituted—similarly, experiences of 'non-perceptible difference' are as a class significantly similar to one another and different from experiences of 'perceptible difference', quite independently of any analysis we may make of the experiences as to the part played by the act of discrimination as distinct from what is discriminated.

Bearing the above considerations in mind, consider now the experience of utility and corresponding welfare. Is it not clear that, no matter how we analyse an individual's utility-comparison of A and B into a pair of utilities and a discrimination between these utilities, the welfare-difference turns on the resultant judgement? The point is a simple one. There is a sense in which it is not unreasonable to say that a judgement, for example, of sameness of a pair of colours is illusory if the failure to judge difference is a failure of discrimination. But to talk of an illusion about welfare is to talk nonsense—or, to put it in another way, an illusion of welfare is as good as reality. Surely an individual is as well off as he thinks he is, and that is all there is to it. A marginal preference is, therefore, an absolute standard for an individual of welfare-difference, and it matters not what analysis we make as to how such preference is supposed to arise from some combination of 'objects presented for discrimination' and 'discrimination between them'. Since there appears to be no way of making such an analysis, this is fortunate, and a marginal preference can be used as standard both for welfare-difference and utility-difference. Whether or not there is some other more ultimate psychological scale for utility can be left as an unanswered and, may be, unanswerable question. The fact remains that we need no such ultimate scale, either for a theory of choice or a theory of welfare, since the point of reference, in both cases, is where indifference ends and preference begins.

Clearly, if this standpoint be granted, the conclusion cannot reasonably be denied that, if an individual's tastes change, this makes no difference to the concept of welfare-difference—a marginal preference will continue to indicate the same welfare difference, and my marginal preference for

A over B now will indicate the same welfare-difference as my marginal preference for C over D yesterday, for E over F in my early childhood. But I believe other persons have preferences and indifferences and I cannot doubt that a marginal preference has just the same significance to them as to me, however difficult it may be for me to get inside their minds in other respects. I do not need to show that marginal preferences are invariant in passing from mind to mind. On the contrary, no such question of invariance can even be raised if my standpoint is adopted.

I must, in conclusion, refer to a question that Rothenberg raises in the last few pages of his article, the question of changes in an individual's utility function. I tacitly ignored this problem in my article as being outside its scope. Clearly, if my tastes change from situation A to situation B, there is no meaning to my preference for A over B, since the requirement for my act of preferring is my contemplation of the alternatives at a given moment when I am what I am. My proposed group-welfare index can do no more than indicate the welfare-differences between alternatives for a group of persons—alternatives, that is to say, that are hypothetical substitutes. It does not even provide a method for comparing my welfare today with my welfare yesterday, unless it can be assumed that my 'ways of preferring' today are the same as my 'ways of preferring' yesterday. That I am able to compare the welfare-difference between A and B yesterday with the welfare-difference between C and D today does not mean that I can compare the welfare of A or B yesterday with the welfare of C or D today.

If, however, my suggestion be adopted that an individual is able, even roughly, to distinguish between positive and negative welfare, and therefore to assign a zero to one of the welfare-positions, even this gap could be bridged. In the above example the preference-intensity for A (over a situation of zero utility) would be determinate—similarly with C. A and C thus become comparable. Group-welfares would also then become comparable, no matter how the groups had changed, whether by changes of taste of the same members or changes of membership. I dare not suggest that so tremendous an edifice should be constructed on so flimsy a foundation, but I can see no other foundation for those judgements that we all do in fact confidently make without waiting for some theory to provide that justification.

UNIVERSITY OF SOUTHAMPTON

THE EFFECT OF TAXATION ON RISK-TAKING

By PAUL STREETEN

ACCORDING to a widespread belief, taxation inevitably discourages enterprise and risk-taking. This paper attempts to collect some considerations relevant to this problem.

Risk-taking presents peculiar difficulties to analysis, which are absent in the case of other productive activities.[1] If it were possible to apply probability distributions to this type of behaviour, one might say that optimal risk-taking is that which is prompted by purely actuarial considerations on the basis of the best expectations available. Deviations from an optimum thus defined would be caused by diminishing marginal utility of income and possibly risk aversion, even in the absence of a tax. But this optimum would conflict with the optimum dictated by utility maximization. There is also a contradiction in assuming perfect competition and equilibrium and also genuine uncertainty about the future. Thus the notion of an optimum amount of risk-taking, deviations from which are caused by a tax, contains several ambiguities.

In the real world it is even less likely that we know very much about the optimum amount of risk-bearing. In some situations (e.g. underdeveloped countries) it may appear fairly obvious that total risk-taking is too small, although even then it is not easy to lump all kinds of risk-taking into one category which is treated much like other productive factors. The shouldering of some types of risks may be desirable, of others undesirable. A shortage of certain types of uncertainty-bearing may be accompanied by an excess of others. Has, then, the total supply of risk-bearing any significance for welfare purposes?

In spite of these difficulties about the determination of the right 'amount of risk-taking', the common argument is that taxes tend to reduce risk-taking and that this is a bad thing, except possibly in booms.

Yet it is not at all obvious that even a very high proportional income-tax[2] would reduce risk-taking. Two sets of questions have to be answered. (1) How are anticipated gains and losses affected by the tax? Though gains will be reduced, losses may also be reduced and thus the net expected

[1] The effects of taxation on uncertainty-bearing (used synonymously with 'risk-bearing') must be distinguished from its effects on investment. Certain types of taxes discourage investment even under conditions of certainty. Cf. E. Cary Brown, 'Business-Income Taxation and Investment Incentives', in *Income, Employment and Public Policy, Essays in Honor of Alvin Hansen.*

[2] Although the following argument is conducted for a proportionate income-tax it applies also to a progressive tax structure, as long as the risk-taking leaves the investor in the same income-tax bracket. There will be no deterrent effect on the score of progressiveness, although, of course, a proportionate tax would yield the same revenue with a lower marginal rate.

yield need not be reduced. (2) How does the uncertainty-bearer react to changes in gains and losses?

Let us take the second question first. *If* the tax reduces expected net income (which is not, as we shall see, necessarily the case), there will be two effects in opposite directions. On the common assumption that the marginal utility of income is greater at the lower income, greater risks, which can be expected to yield greater incomes, will be taken in order to get income. On the other hand, since the reward for marginal risk-taking has been reduced, fewer risks will be taken. The net outcome will depend upon whether the increased need to get income is sufficient or insufficient to offset the discouragement of a lower premium.[1]

It might seem as if we could say that if the marginal utility of income is falling rapidly, more risk will be taken after a tax is imposed. But this may be misleading, for what matters is the rate of fall in the marginal utility of income *relative* to risk-taking. Now the marginal disutility of risk-taking may be increasing equally or even more rapidly as risk increases. Moreover, there may be special 'gambler effects' which are different for different persons.

One way of interpreting the situation is that suggested by Milton Friedman and L. J. Savage. From some, though slender, evidence of choices involving uncertainty[2] they conclude that, over a range, the marginal utility of income is rising. Before the tax is imposed, the individual receives an income OM (Fig. 1). At OM the marginal utility of gains (CD) is smaller than the marginal utility of equivalent losses (AB) since the marginal utility of income is diminishing. A tax reduces net income to ON. Now the marginal utility of income is increasing. Therefore the marginal utility of expected gains ($C'D'$) is greater than the marginal utility of equivalent losses ($A'B'$). It is true that, if the position prior to the imposition of the tax had also been in the range of increasing marginal utility, both the income effect and the substitution effect would have tended to reduce the amount of risk taken. At the lower income both the

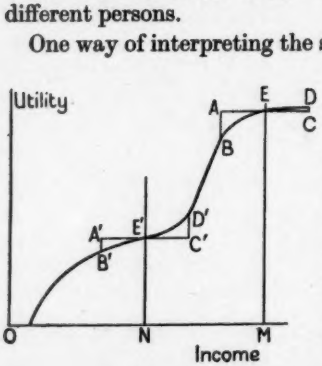

FIG. 1

[1] In equilibrium $\frac{du}{dy}\cdot\frac{dy}{dr}=\frac{v}{dr}$, where u is the utility of income, y income, v is the disutility of risk-taking and r the amount of risk-taking. A tax will increase du/dy, if the marginal utility of income diminishes with income, but it will diminish dy/dr, so that the net change in risk-taking is uncertain.

[2] e.g., the preparedness of most people to gamble and to insure and the fact that there is more than one prize in lotteries. Cf. 'The Utility Analysis of Choices Involving Risk', *Journal of Political Economy*, Aug. 1948.

marginal utility of income and the marginal reward for risk-taking would have been less. But we should not expect to find many people in this range. Since they are keen gamblers, prepared to accept unfavourable odds, they will tend to be pushed either into the lower range, or, if they are lucky, into the higher range, of diminishing marginal utility. By a process of natural selection the middle range will not survive. In most cases the position before the imposition of the tax is one of diminishing marginal utility and we therefore cannot be sure about the direction of the income effect. The utility of marginal income may be greater or less after tax. But if the individual finds himself in the range of increasing marginal utility, there will now be an additional 'gambling' effect. He will be prepared to accept worse than mathematically fair chances (if he wants to maximize his utility) and his risk-taking will increase. The nearer the point *N* moves to the left towards the lower range of diminishing marginal utility, the less likely will the individual be to increase his risks. But a tax which leaves him in the range of increasing marginal utility will have a strong effect towards greater risk-taking.[1]

Friedman and Savage offer a sociological interpretation of the wiggly shape of the income-utility curve (two segments of diminishing marginal utility and one segment of increasing marginal utility). The two concave segments correspond to different social and economic levels (determined by the income of other members of that class) and the convex segment to the transition between the two levels. An increase in income that leaves the individual in the same class is subject to diminishing marginal utility; but one that shifts him into a new class, to increasing marginal utility. Hence fair chances will be refused if they cannot be expected to yield the opportunities of a new status, whereas even unfair chances which might lead to such a rise, but are more likely to throw him back into a lower section of his own class, are accepted.

In this situation we would expect uncertainty-bearing to be particularly stimulated, at least for a time, whenever a tax which is not so high as to cause despair pushes income below the level appropriate to the class to which people are used to belong (from *OM* to *ON* in Fig. 1). They will hazard a good deal more now than before in order to attempt, however uncertain the outcome, and in spite of great dangers, to maintain their social position.

But it is sufficient to assume diminishing marginal utility of income

[1] Daniel Bernoulli already envisaged the possibility of increasing marginal utility of income over a range and gave the following example as an exception to the principle of diminishing marginal utility: 'Thus in most cases the same gain will mean more to a poor than to a rich man; nevertheless, to a prisoner who possesses 2,000 ducats and who needs another 2,000 to regain his liberty a gain of 2,000 will mean more than to a poorer man.' The fact that interpersonal comparison is involved is irrelevant for our purposes. He could have equally well considered the same man in different circumstances. Cf. *Versuch einer*

relative to risk-taking throughout, in order to show that the effect of a tax on risk-taking may be in either direction. Income effects will be strong, the more rapid the fall in the marginal utility of income relative to the marginal disutility of risk-taking.

In Fig. 2 we measure on the ordinate risk in the sense of expected losses, expressed as a percentage of the amount invested. They are computed as

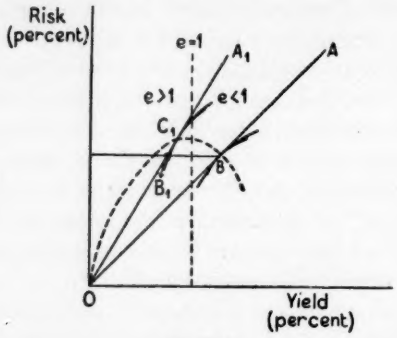

FIG. 2.

the actuarial value of a subjective probability distribution,[1] but the argument could be adapted to either the most probable losses or focus losses.[2] On the abscissa we trace expected yield as the mathematical expectation of all expected outcomes (positive and negative), expressed as percentage values of the amount invested, analogously to Keynes's marginal efficiency of capital. (The analysis could be adapted to most probable yield or focus gains and it could be conducted in terms of absolute, rather than percentage, values.) Risk and yield are assumed to be the only relevant variables. On the assumptions of diminishing marginal utility of income and increasing marginal disutility of risk-bearing, the indifference curves will be positively sloped at a diminishing rate and getting flatter as we move to the right horizontally, and as we move upward vertically.

The market opportunity curve shows the views of the investor about the possibilities open to him of various optimum risk and yield combinations. It ought to be defined as a notional curve since no objective risks and yields exist. It therefore depends partly on market opportunities, but also

neuen Theorie der Wertbestimmung von Glückefällen, translated and edited by A. Pringsheim, Leipzig, 1896 (*Specimen Theoriae novae de Mensura Sortis*).

[1] The analysis follows that of Evsey D. Domar and Richard A. Musgrave, 'Proportional Income Taxation and Risk Taking', *Quarterly Journal of Economics*, May 1944. As the authors point out, to assume wealth as given implies that the indifference map shifts with any change in wealth. An alternative would be to measure absolute amounts on the axes, in which case a change in wealth would shift the opportunity curve.

[2] G. L. S. Shackle, *Expectation in Economics*, ch. ii.

on the optimism or pessimism of the investor and upon his beliefs about the opportunity lines of others.

It is assumed that a no-risk-no-income position is open to the investor, so that the curve starts at the origin.[1]

The shape of the market opportunity curve may be concave or convex to the abscissa or straight over a range, until it bends backwards. Convexity is suggested by the possibility that an investor who is taking many risks, say by extreme illiquidity, can, at first, reduce risks faster than income. Even a small sacrifice of income may prevent the necessity of forced sales in unfavourable conditions.[2] When risks are already greatly reduced, large sacrifices in yield are required to reduce risks further.

Concavity, on the other hand, is suggested by the fact that the chances of getting yield by increasing risk-taking are greater (within limits) the more risks are taken, for two reasons. First, because *others* may be unduly pessimistic about investment opportunities and their opportunity curves may be unduly convex. Second, because risk aversion of others (the concavity of their indifference curves) may reduce the number of people in the field who compete for the rewards of risk-taking. For either of these reasons or a combination of them, an investor who is less pessimistic and more daring than the average will be faced with a concave curve. If there is to be stability it would have to be less concave than his indifference curves and it could be expected to become convex at some point.

Assuming, to start with, a straight opportunity line OA, a tax will shift each point on it to the left. We assume that there are no loss offset provisions, so that for any given risk positive gains only are reduced proportionately and hence yield is reduced more than in proportion to the tax. Only if there was no chance of losses would yield be reduced in proportion to the tax.[3] If the investor undertook the same venture as before he would now be at B_1. But this is not an equilibrium point for him and he will shift until he reaches C_1, which may lie above or below B_1, i.e. he may now take more or less risks. The points of tangency of various tax rates will describe a curve OC_1B. At first, as the risk price of income is raised, more risk is taken to get income. In this range the elasticity of demand for

[1] Since the analysis is carried out in terms of money, it is possible that the market opportunity line should start at a positive point on the yield axis. This would be the case if inflation in the future is certain. The fact that money is the only asset expected to lose value is expressed by the fact that a riskless income (i.e. appreciation) is forgone if money is held.

[2] The possibility of borrowing will reduce risks to some extent. But it can only delay the point at which the risk price of yield rises sharply. Lenders' risk will render the terms of borrowing increasingly severe, and borrowers' risk will, even at constant terms, make risks rise faster than yield. For yield falling short of the fixed interest rate over a time would mean bankruptcy.

[3] The fraction by which yield is reduced by the tax is $(1+r/y)t$, where r is risk, y is yield, and t the tax rate (cf. Domar and Musgrave, loc. cit.). Therefore, if r/y increases as we move up the opportunity curve, yield will be reduced by a greater proportion, the higher we are on the OA_1 curve, whereas if r/y is falling, yield will be reduced by a smaller proportion.

income in terms of risk-taking is less than unity. At the top of the hill the elasticity of demand is unity and finally less risk will be taken, i.e. the elasticity of demand is greater than unity. If the demand for yield in terms of risk-taking is completely inelastic, more and more risks will be taken as the tax rate increases. OC_1B will be a vertical line.

If the opportunity curve is not a straight line but convex to the abscissa, i.e. if the risk price of income increases, the fraction by which yield is

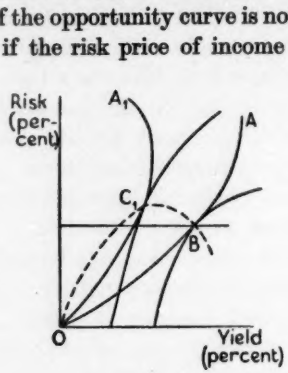

FIG. 3a.

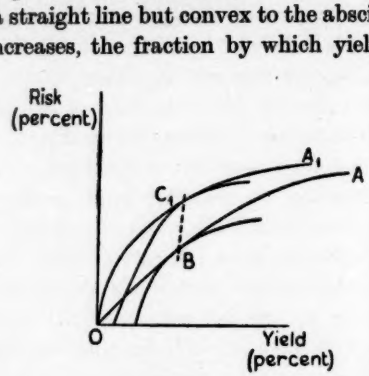

FIG. 3b.

reduced by a given tax rate increases as we move up the opportunity line. This will reinforce the substitution effect towards reduced risk-taking (see Fig. 3*a*). If, on the other hand, the opportunity line is concave to the abscissa, i.e. the risk price of income falls, a given tax rate will shift every point on the line less to the left the higher up we move and this will reinforce the income effect towards increased risk-taking (see Fig. 3 *b*).

We can conclude that even without any loss-offset provisions risk-taking may be increased after a tax. It will increase by more, the stronger the income effect and the lower the rate at which risks increase as yield is increased.[1]

The foregoing analysis has serious limitations. I shall not, however, discuss the difficulties arising from the assumption that investment is perfectly divisible and hence adjustments always possible along continuous opportunity lines.

Another limitation of the analysis is that decisions are assumed to depend upon two variables only, expected yield and risk, or focus gain and focus loss. This is clearly unrealistic.[2] Qualities of the probability dis-

[1] Many authors believe that a proportionate tax necessarily discriminates against risky investment. See, e.g., E. F. Schumacher, in *The Economics of Full Employment* (Oxford Institute of Statistics), p. 94; J. R. Hicks, U. K. Hicks, L. Rostas, *The Taxation of War Wealth*, p. 193; Emil Schram, 'Taxation and Venture Capital', *The Annals of the American Academy of Political and Social Science*, Nov. 1949, p. 88.

[2] In addition there is the general difficulty of using any kind of certainty equivalent. It assumes away some of the essential characteristics of behaviour in conditions of genuine uncertainty.

tribution which are decisive for decision-making are concealed. These qualities may be differently affected by different taxes. In particular progressive taxation will make a difference according to the shape, not merely the mathematical value, of the probability distribution, and in particular according to how the upper end of the distribution is valued.

The assumption that expected yield and expected risk are the only relevant variables implies that the investor is indifferent between ventures with different probability distributions but the same yield and risk. This is obviously untrue. Thus, for a given yield, greater risks might be accepted if there is a small probability of large gains even though the probability of small gains is small, whilst these risks would not be taken if the probability of small gains is large but there is no chance of a large gain. A greater dispersion of the expected gains may be experienced as pleasant or as unpleasant. If the lure of large gains is strong a progressive tax would have a more deterrent effect than a proportional tax which would reduce yield by the same proportion.

In addition to yield and risk, the prospect of obtaining less than a minimum rate of return might be very frightening;[1] or the chance, though slight, of securing more than a minimum gain may be particularly attractive. Or again, the possibility of suffering losses above a certain amount may loom particularly large.[2]

The importance of minimum yield or of maximum loss is ignored by the use of two variables only, yield and risk. On the other hand, Professor Shackle's focus values ignore other relevant properties of the distribution of potential surprise.

A particular case of the possible importance of getting a minimum gain is the situation of small firms. It is true that some might, in order to avoid bankruptcy, refrain from taking some risks when their income is reduced through taxation. But, on the other hand, they might be forced to accept greater risks in order to secure at least a chance of survival. A firm with a large income may well prefer to forgo some income in order to consolidate and strengthen its financial position. But one with a small income might only be able to avoid certain death by increasing its risks, even although the gamble might result in disastrous losses. This may be due to the shape

[1] By selecting risk as a variable we assume that an expected return smaller than zero is the deterrent. But there may also be a return less than a certain positive amount which the investor particularly fears. The analysis would have to be modified accordingly.

[2] This is an alternative way of interpreting the Friedman–Savage case. The same behaviour can be rationalized either by a peculiar shape of the income-utility curve (see Fig. 1) or as a peculiar psychological attitude to certain types of uncertainty. The Friedman–Savage formulation maintains the principle that expected utility is maximized and sacrifices the principle of continuously diminishing marginal utility. Alternatively, continuously diminishing marginal utility can be maintained, but the simple principle of utility maximization is sacrificed by the introduction of the pleasure of gambling, the fear of large, though uncertain losses, &c.

of the opportunity curve: At high incomes a small reduction in income may 'buy' a large reduction in risks, whereas at small incomes a large reduction in income 'buys' only a small reduction in risks. Or it may be due to the slope of the indifference curves, or a combination of the two.

The position of a person far below the subsistence minimum is similar. If he has no other source of income, he may be forced to invest all his meagre assets. He cannot afford to be liquid, i.e. cannot afford *not* to take risks.[1] The position is analogous to that of a person trying to maintain a customary level by increasing his risks. A man with nothing in the world except one shilling might well jump at a 1:100 chance of getting 80 shillings or losing his shilling. The position of a customary level, below which the marginal utility of income falls off very slowly, but above which it falls rapidly is illustrated in Fig. 4. Increased tax rates without loss offsets will increase risk-taking and have precisely the same effect as taxes with full loss offset provisions and possibilities.

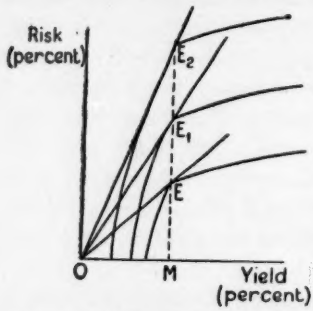

FIG. 4. Discontinuity at customary or minimum required income level *OM*.

At higher tax rates, more risk is taken.

Below *OM*: 1. ind. curves steep; 2. slopes fall slowly along a constant risk line.
Above *OM*: 1. curves are flat; 2. slopes fall rapidly along a constant risk line.

Next the question of loss offsets has to be discussed. Persons or firms who can deduct losses suffered in one line of investment against income earned in another line or against past income because (1) they succeed in earning such income and (2) the tax laws allow them to do so, are, of course, capable of shifting some of the losses to the State. The State thus becomes, in the fullest sense, a partner not only in gains but also in losses.

A tax with full loss offset leaves the ratio of yield to risk unaltered. In Fig. 5*a*, B_1B/OB is the tax rate, reducing both risk and yield by the same proportion. If the opportunity line is straight, the equilibrium point *B* must remain the same as before. The effective risk price of income has not been altered.

If the opportunity line is convex to the abscissa, more or less private

[1] The figures produced by Mr. W. L. Crum's investigation *Corporate Size and Earning Power* (Harvard, 1939) appear to confirm the thesis that amongst larger entrepreneurs those who earned profits had a lower profit rate, but had it more frequently; this would suggest that the advantages of larger-scale production and credit facilities are partly enjoyed by *reducing* risks. Though those amongst smaller entrepreneurs who did earn profits earned a higher rate than their larger competitors, the average of all, including those who incurred losses, is smaller. This suggests that smaller firms shouldered greater risks. These conclusions are drawn tentatively by J. Steindl, 'Capitalist Enterprise and Risk', *Oxford Economic Papers*, No. 7, March 1945.

risk will be taken after the tax (C_1 may be above or below B in Fig. 5*b*). There will be only an income effect and no primary substitution effect, though a secondary substitution effect will arise from the fact that the risk price of yield increases as we move along the opportunity line.

But even although private risk may be reduced because the Treasury shares now in losses, this underwriting of private risk must increase total

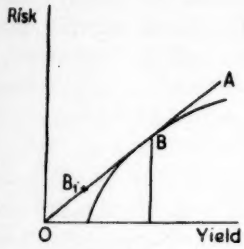

FIG. 5*a*.

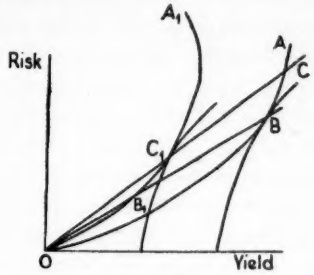

FIG. 5*b*.

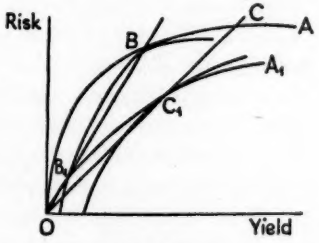

FIG. 5*c*.

risk. The new position of the investor before tax is deducted is not C_1 but C, the point where the line OC_1 extended cuts the curve OA. Total risk must increase. There is the difficulty of defining total risk. The opportunity line shows merely the evaluation of the opportunities by the investor. Moreover, it is doubtful whether the risks of several investors can be summed up into a total shouldered by the investors and the Treasury together. Nevertheless, the risk as estimated subjectively before tax is deducted (at C) must be the same as or greater than in the pre-tax situation (B). Full loss offset provisions are a clear encouragement to the total risks incurred.

If the opportunity line is concave, a tax will benefit the investor, enabling him to reach a higher indifference curve (Fig. 5*c*). Highly risky ventures will now become profitable. Private risk may be increased or reduced and total risk must be increased.

The question of what losses can be set off is, of course, a matter of law.

They need not take the form of capital losses. They may take the form of a lower taxable income, because fixed interest charges, depreciation, or other commitments which must be met in the event of failure are set off against other income. Any of these may be charged against other taxable income.

It is clear that loss offset provisions discriminate in favour of those with large incomes, the established firm with past earnings, and the firm that is carrying on a variety of operations, and against the small, the new, and the undiversified firm.[1] Moreover, research and experiments in new lines with new products or new productive processes are often known to involve initial losses. A tax which allows for loss offsets will encourage such experiments which can be financed partly at the State's expense. Since new entrants will not be capable of similar experiments, having no income against which to offset the losses, the established concerns will be able to forge ahead into new fields and add further to their power and income.[2]

The likelihood that income is available for loss offsets is greater if the law provides for carry-forward and carry-back of losses. Small and new firms could be put on an equal footing with large and established firms by State subsidies in case of losses where no other income is available.

There are, of course, other methods of reducing or avoiding the deterrent effects of taxation besides loss offset provisions.

Mr. Kalecki's scheme[3] of a modified income-tax is designed to leave the expected profit rate on real investment unaffected by the tax. Income-tax is charged on gross income, i.e. before deduction of wear and tear, but all investment in fixed capital, whether replacement or expansion, is deducted from the taxable amount at the time when it is made. If it exceeds the taxable income the excess is carried over for deduction in subsequent years.[4] The rate of profit on new investment would thus not be affected. Assume that an investor requires a rate of return of 10 per cent. to induce him to invest £100. If his investment yields at least £10 per annum he will embark on it. Now an annual tax of, say, 10 per cent. will reduce his

[1] There is evidence that in the U.S.A. the discouragement (if any) to individual risk-taking through high tax rates may have been largely offset by an increase in the availability of undistributed profits to large corporations. 'Although corporate tax liabilities rose from $1.5 billion in 1939 to $13.0 billion in 1949, corporate profits after tax stood at $20.9 billion in the latter year as compared with $5 billion in the former' (E. Gordon Keith, 'Repercussions of the Tax System on Business', in K. E. Pool (ed.), *Fiscal Policies and the American Economy*, 1951, p. 347). This may be partly due to the income creating effects of tax expenditure.

[2] The losses must not, however, be expected to last longer than the period allowed for carry-forward by the tax law.

[3] M. Kalecki, 'Three Ways to Full Employment', *The Economics of Full Employment* (Oxford Institute of Statistics), pp. 45–46.

[4] To this carry-forward allowance an appropriate interest rate may be added in order to put it on an equal footing with allowances to firms with sufficiently large incomes to receive the whole deduction at the time of investment expenditure. There is, however, still the danger that income does not cover the amount which could be written off and smaller firms will still be worse off than larger ones.

expected annual gross profit to £9 and the investment would not be carried out. But according to the Kalecki scheme by investing £100 he would get a £10 reduction in tax. The cost of investment would therefore be only £90 and the expected profit net of tax (£9 out of £90) is still 10 per cent.

The main difficulty of such a scheme would be the necessity to distinguish between real and purely financial, and between new and second-hand, investments. One would not want to exclude all financial investments because this would be an undesirable discrimination against joint-stock companies. On the other hand, one would want to exclude financial investments which are merely transfers of paper assets. A way round this difficulty may be found by a careful classification by an authoritative body of securities which would qualify for exemption.

The purchase of second-hand equipment, though not new investment itself, may give rise to it indirectly. Since smaller firms depend on it more than larger ones, provisions which would disqualify it from tax exemption would discriminate against the former.[1]

It could be argued that such a scheme would go too far the other way. Thus the right of a firm to write off against taxable profits the whole cost (minus scrap value) of its plant,[2] even though it may turn out a complete failure, positively encourages risk-taking. It goes beyond full loss offset, for, in the event of success, no tax on capital appreciation would have to be paid.

A weaker proposal would be to *accelerate* depreciation on new investment in the tax-exemption provisions. Kalecki's proposal is the extreme case of allowing full depreciation in the same year in which the asset is bought. If there is uncertainty about whether sufficient income will be earned in the more distant future to allow for deduction of depreciation allowance, this uncertainty can be removed inasmuch as present income and income in the near future is more certain. The allowance would also be greater than one spread over the life of the asset, by the interest payments on the initial outlay.

Many firms buy new equipment only if it is expected to pay for itself in 5 years or less. If depreciation allowance can be spread over these 5 years, instead of over the physical life of the asset, such investments would be encouraged. If it is desired not to allow the full benefit of this concession to larger firms, limits could be set to depreciation allowances deductible from taxable income.

The argument has so far been conducted on highly abstract lines. The drawback of such abstractions is that they often single out easily manageable variables and neglect the really important features of complex actual

[1] Cf. E. Gordon Keith, op. cit., p. 354.

[2] Under the existing income-tax law a firm can write off the full value of abandoned plant minus scrap value against taxable profits. Apart from the rate of interest, this has the same effect.

situations. Several characteristics may be mentioned which appear to be amongst those which play an important part in actual situations, and which would require substantial modifications of the preceding analysis:

1. Incentives depend largely on social conventions.
2. The existence of joint-stock companies requires modification of the analysis.
3. There may be risk preference rather than aversion at the margin.
4. The attitude to risk-taking will vary with business conditions.
5. A part of what seems to be a return on uncertainty-bearing is monopoly revenue.
6. The existence of opportunity costs must be taken into consideration.
7. Profit expectations are affected by the imposition and expenditure of a tax.

1. Henry Simons wrote in his *Federal Tax Reform*, p. 146:

> One may venture the notion that the adverse effect of progression upon incentives also depends more largely upon prevailing social attitudes than on anything amenable to investigations by devices of the professional economists. . . . Top surtaxes which are widely regarded as either reasonable or permanent, and accepted as established rules of the game, need not much impair the productional activity of a competitive enterprise system.

The expected rewards for uncertainty-bearing are partly a function of what others similarly situated expect to get and what similar ventures have yielded in the past. In so far as this is important, the problem becomes largely one of the transition period, until the new and lower net returns have come to be accepted as appropriate.

2. The existence of limited liability introduces special features. First, by enabling the investor to spread his risk over a series of ventures, it makes risks smaller than what they would be if he had to put all his eggs into one basket. The possible income variations are reduced and hence the impact of progressive taxation is less deterrent than it would be otherwise.

Second, in corporations where management and often control are separated from ownership a reduction in dividend payment to shareholders need not affect the willingness of those responsible for decisions to take risks. Their main interest is the growth of the company. A personal income-tax may, of course, induce shareholders to sell shares and switch to safer assets. This would make it more difficult to recruit financial resources for the enterprise. But this effect is mitigated by the possibility of internal financing out of retained profits.

A progressive income-tax will, however, discriminate against high yield assets. Their yield may be high because they are expected to depreciate. Assets which are expected to appreciate and have therefore a low yield

will, on the other hand, be taxed less. Such discrimination could be avoided if the tax were imposed not on yield but on the capital value of the asset.

3. It is not at all clear that risk-taking is a real cost rather than a positive utility. It is doubtful whether the concept of a 'margin of risk-taking', which presupposes a homogeneous factor of production, makes sense. Even if it did, a producer's surplus or rent might be enjoyed at the margin. There can be no doubt that many types of uncertainty lend excitement, romance, and sometimes dignity to a person's activities. It is not average temperament that is relevant here. If there are people who like uncertainty it will be they who will tend to be recruited into the risk-taking classes, provided, of course, they possess the necessary resources.

If it is replied that the very fact that we do not expose all our resources to uncertainty suggests that uncertainty-bearing must be a disutility at the margin, the answer is: the limitation to risk-exposure may be due to diminishing marginal enjoyment rather than increasing marginal nuisance. Resources used in ways other than risk-exposure yield even more fun, after a point, than resources expended on gambling.

The question of the psychological attitude to uncertainty-bearing (the shape and position of the indifference curves) must be distinguished from the question of optimistic, pessimistic, or correct forecasts (the shape and position of the opportunity curve). The latter refer to expectations, the former to the psychological weights attached to these expectations and the investor's attitude towards them. Keen uncertainty bearers may be also constitutionally over-optimistic. In this case their net average returns over time will tend to be negative (unless monopoly elements enter). But the gains of a few may be sufficient to attract ever new victims.[1] If, on the other hand, their estimates are over-pessimistic they will tend to earn positive returns (though not as high as if there were also risk aversion).

[1] 'Historically examined, the oil industry, and particularly the production branch, has displayed more than a normal amount of spirit of the game and adventure. The search for what is properly termed "black gold" partakes of the nature of the hunt for precious stones and metals. It is characteristically uncertain, even with modern development techniques, offers occasional great rewards for a few, and a mediocre competence, or economic loss, for many. This is believed to represent the general situation, even though the larger and better financed operators may be able to lessen the risks in some degree through diversification. It is generally agreed that more gold has been spent in seeking this valuable metal than has ever come out of the mines. Likewise, it is the belief of the industry that much more money has been put into the production branch of the petroleum business than has been realized. This statement must be interpreted in a statistical sense, since common observation, in the leading oil production centres, demonstrates that certain individuals make a great deal of money from the search for and extraction of, crude petroleum. As with the Irish Sweepstakes, the appeal lies in the chance of winning, not in the mathematically much greater chance of losing. And just as many folk think of the Sweepstakes in terms of the newspaper accounts of families made deliriously happy, so the optimism of mankind causes the investor, or speculator, in crude oil possibilities to remember the instances of those few who "without a shirt to their backs", sold out for ten million' (Ronald B. Shuman, *The Petroleum Industry*, 1940, pp. 26–27, quoted by Harold M. Groves, *Postwar Taxation and Economic Progress*, New York, 1946, pp. 10–11).

Rash action in the face of pessimistic expectations tends to have the same effect as more cautious action in the face of optimistic expectations. A tax can reduce expected net returns, but it may do nothing to curb those who delight in dangerous activity.

One objective factor would suggest that pessimistic estimates would be correct, i.e. that unfavourable outcomes have more disastrous results than favourable outcomes have cheerful results.[1] This would suggest convexity of the opportunity curve. Business losses tend to be cumulative in a manner in which gains are not. Losses may cut off a person from credit, may weaken his bargaining position, destroy goodwill, &c. Though there is some tendency for gains to be cumulative too, this is often not so pronounced. But in spite of this objective tendency, the temperament of an investor may be such that, although correctly estimating the situation, he derives greater pleasure from expected (though non-cumulative) gains than from the prospect of avoiding possibly cumulative losses.

4. The asymmetry of the objective factors making losses more cumulative than gains is more important in bad than in good times. Obviously, risk-taking varies with business conditions. During a depression, when bankruptcy is an imminent danger, possible future profits accruing from a venture are discounted at a very high rate. In such conditions even a small tax may be very discouraging. Moreover, it is then less likely that losses, if incurred, will have profits, otherwise taxable, to be set off against. In times of easy profits and general prosperity, on the other hand, losses are feared less and, if suffered, can be set off against gains, and even high taxes will do little to reduce risk-taking. Thus in times when full employment is ensured there is less danger of discouraging risk-taking through taxation than in times when cyclical movements are anticipated.

But although the objective factors increase the chances of unfavourable outcomes when general demand declines, and reduce them when it increases, there is, as we have seen, at least one important reason why *willingness to bear risks* can be expected to move in the opposite direction. When demand is expected to increase, losses will be a stronger deterrent than when it is expected to decline. One important objective of most firms is to stay in business. In conditions of declining demand there is, for some at least, a chance of achieving this objective by accepting certain risks, whereas in conditions of constant or increasing demand the objective is more likely to be achieved by shunning these risks.

5. Perhaps the most important qualification is that returns for risk-bearing cannot be disentangled from monopoly returns due to the limitation of resources available for risk exposure. The supply of capital to any

[1] Cf. Professor A. G. Hart's principle of the linkage of risks. This point has been stressed also by Professor G. L. S. Shackle.

given risk-taker is not infinitely elastic. He cannot and he will not borrow at a constant rate of interest as much as he would want to in the absence of uncertainty and simply in the light of his profit expectations. Returns will therefore include an element of monopoly revenue which is the result of the obstacle put up by uncertainty about the future. To the extent to which rent is earned at the margin, a tax which falls on this rent element can have no effect on the supply of uncertainty-bearing. The monopoly revenue arising from limitations in the supply of capital will tend to be greater for larger firms[1] and this would suggest that discriminating taxation may be advisable. But it will be difficult to disentangle in any situation the revenue which is due to this monopoly element from high returns on capital due to strong risk aversion or due to over-pessimistic estimates about the future.[2]

6. So far, opportunity costs, i.e. the income sacrificed by not investing in alternative lines, have been ignored. But it is clear that a shift to the left of the opportunity curve will have different effects according to whether the investor's pessimism about risky ventures has increased, so that he now believes that for any given risk expected yield is smaller, or whether the leftward shift is due to the imposition of a general proportionate income-tax. In the latter case the returns from alternative lines have been reduced too. Since the indifference map was drawn up on the assumption that other investment opportunities are given, the whole map ought to be redrawn. Alternative earning possibilities are reduced too, and this must weaken the substitution effect and strengthen the income effect.[3]

If, for example, it is the proportionate difference between a risky and a safer investment yield which determines the investor's preference for the one over the other, this difference will be unaffected by a proportionate income-tax and he will not shift away. Both reward and opportunity cost of uncertainty-bearing have been reduced.[4]

[1] If their economies of scale and cheaper credit facilities are not offset by the difficulty to secure a market for their products. Cf. J. Steindl, 'Capitalist Enterprise and Risk', *Oxford Economic Papers*, No. 7, March 1945.

[2] A tax may, of course, reduce the capital *available* for risk-bearing without reducing the *willingness* to bear risks in the schedule sense. The effect on capital availability, or risk-bearing *capacity*, is not examined in this paper.

[3] It is, however, true that the risks arising from reduced liquidity will be increased by a reduction in expected net earnings. There is no advantage to offset this type of increased risk. But as has been argued above, reduced liquidity will not necessarily lead to a reduction in risk-taking. If the motive to stay in business is strong, and if this can only be done through reducing liquid reserves, risk-taking will be increased.

[4] There are, of course, other alternatives besides shifts to safer investments whose relative attractiveness may have increased, e.g. consumption or hoarding whose opportunity costs are zero before and after tax. But individuals do not normally shift into consumption when their risky incomes are taxed, except, possibly, at very high rates. The relatively greater attractiveness of hoarding is discussed below.

Even if the tax on risky income were discriminating so that safe income is exempted, the substitution would be limited by the reduced yield of safe income resulting from a general shift away from risky investments.

To this argument it is frequently replied that the decisive question is not the proportion between risky and safe income, but the proportion between risky income and the cost of the capital asset and thus the *absolute* difference between risky and safe investment that determines the willingness to expose wealth to uncertainty. A reduction in this proportion of income to capital value (amount of difference between uncertain and certain income) means that amortization will now take longer.[1] This is said to reduce the attractiveness of the uncertain investment. In the extreme case of a 100 per cent. income-tax the safe investment would still be valuable as purchasing power, whereas the risky investment can only involve a net loss.

This argument cannot refer to a longer period of *waiting* for a return of the outlay, for the time in which the initial outlay is recouped (if interest plus amortization is paid on the safe loan and both are subjected to tax) is prolonged even further in the safe line. It must refer to the possibilities of losses of capital, as well as income. This is obvious, because an investor can always cut short the time of exposing an asset to uncertainty by selling it before full amortization. But this may involve a loss, not only because the prospects of the asset may have deteriorated but also because buyers, rightly or wrongly, do not share the optimistic views of the seller about its prospects. Thus there is clearly a chance of capital losses which can be escaped by playing safe.

But where there is a danger of capital losses there is also a hope for capital gains. If the latter are untaxed or taxed at a relatively low rate, there is no reason to assume that the fear of losses will always be stronger than the hope for gains. It is well known that very high rates of income taxation can be an incentive to risky investment. Returns can be taken in the form of capital gains, even though dividends are taxed at a prohibitive rate. In order to be able to predict what would happen, the psychological factors discussed above would have to be examined first. As far as the subjective factors are concerned the result is quite uncertain.

Objectively, a reduction in possible gains which leaves losses unaffected will tend, as we have seen, to impair the credit standing of the individual and his enterprise. The margin of net expected profit will be narrowed, both as a result of lenders' risk which will have to be compensated by a higher rate of interest, and also by borrower's risk which will make the investor more reluctant to borrow at any given rate of interest if the absolute excess of expected (but uncertain) returns over costs is reduced by a tax. In the extreme cases he may be cut off completely from capital supply. This argument suggests that such a tax will discriminate against

[1] Cf. Schumacher, loc. cit., p. 94. This is, of course, not true if amortization is exempted from taxable income unless net income in any future year during the life of the asset is expected to fall short of depreciation allowance so that it will be lost for lack of income.

those who have to borrow funds to undertake enterprise and in favour of the established and/or large firm which can use its own funds.[1]

7. Finally, it must be remembered that expectations about profits may be affected by the imposition of the tax. Changes in taxation are often taken to be indicators of government expectations about future business conditions. Thus increases in tax rates are often connected with the idea that the government expects inflation. If there is strong evidence that risky assets will not depreciate in the future or are bound to appreciate, one could argue that the investment in shares thereby becomes 'safe'. A shift into money would then be a risky venture. In any case, the more widespread the expectation of inflationary conditions, the more likely it is that even a very high income-tax would not discourage what is traditionally called 'risky' investment. High tax-rates coupled with easy profits will make the rate of discount of profits in the more uncertain future low (especially if tax-rates are anticipated to be reduced later), and this will tend to encourage embarkation on speculative ventures.

The next step would be to examine the effects of taxation on incentive in the aggregate. One would have to compare two specified situations and the use that is made of the tax receipts by the government. Thus if the tax is used to create a budget surplus, the incentive effect of the tax would have to be compared with the effect on incentive of the conditions without the tax. The situation in the latter case may be more inflationary or, with a sufficient stimulus to production through tax relief, discouragement to dissaving and large expenses accounts, possibly less. If it is more inflationary, the type of inflation to be envisaged would have to be examined. Suppressed inflation is likely to reduce the attractiveness of marginal income since its spending power is restricted. Open inflation is likely to overstimulate risk taking.

If the tax is used to repay debt, future taxes will be lower than they would have been if the national debt had not been reduced.

If, on the other hand, the proceeds of a tax are used for social transfer payments one would have to take into account the incentive effects of income-increasing or income-creating subsidies, as well as of the taxes. Their impact on average and marginal income would have to be examined. An inquiry into the incentive effects of a tax which neglects the direct and indirect effects of government expenditure on the level of activity and on incentives is only of limited interest.

[1] We have seen that loss offset provisions have the same tendency.

SCARCITY, MARXISM, AND GOSPLAN

By P. J. D. WILES

1. THIS article is a survey of the incompatibility of Marxism with 'scarcity' economics, i.e. the proposition that the essence of economics is choice, or the allocation of scarce resources between competing ends. The word 'scarcity' will be used here as shorthand for this concept. It is argued that this problem has never been and is not now understood by Marxists, in or outside U.S.S.R. We deal first with the Founding Fathers, then with monetary and accounting theory, price and output formation, and planning. Sections 8 to 17 examine the direct statements of Soviet economists on this topic, and finally it is suggested that 'choicelessness' is not really a great drawback at all. The writer's own opinion on 'scarcity' should perhaps be stated. I believe it to be a very important problem, but on a par in importance with a few others such as unemployment (which is by definition the opposite of scarcity), the growth and distribution of the national income, technical progress, inflation; that while economics is certainly not to be defined as the science of choice, no one is a good economist who does not always take it fully into account.

It must also be insisted that 'scarcity' does not only arise in an economy with consumers' sovereignty, as is often thought. It is true that 'scarcity' was invented by capitalist economists who believed in consumer's sovereignty, and attacked centralized planning for neglecting both. But if all consumers are rationed in whatever way the State pleases, and the principal interest of the State is industrialization, relocation of industry, militarization, or what not, the State faces the identical problem of scarcity that in the *laissez-faire* economy was previously faced by consumers. Resources are still scarce, and ends still compete. The only difference is that the ends have changed. All economies face a scarcity problem, and the strict linkage of it to consumer's sovereignty is due to historical accident.

Nor should free consumer's *choice* be confused with the general rationality of the economic system. So long as he is unrationed the consumer can exercise choice in spending his own income. But if the producer is not subject to the pull of demand and profit the consumer is in no way sovereign and the general problem of choice is no nearer to solution (unless of course the State is imposing some conscious, centrally planned solution) than when there was rationing. In U.S.S.R., where we are assured that production is not for profit, the abolition of rationing has no effect on relative quantities of output or on the rationality of the whole economy. There is either a centrally imposed rationality or none at all.

2. We should never underestimate the Marxist orthodoxy and Marxist continuity of Soviet thought. Even in such a field as economic planning, about which Marx consciously said nothing, the tradition imposes certain limitations and encourages certain attitudes. We begin therefore with Marx himself. 'Scarcity' arose out of the subjective theory of value. Although the Menger–Jevons revolution occurred more than 10 years before his death (1883), Marx seems never to have written one word on the subjective theory of value itself: striking proof of his self-confidence and self-isolation. *A fortiori*, then, he neglected scarcity. In Engels's letters I have been able to trace five references to subjectivist economics and none to scarcity.[1] Though he is extremely sure the theory is wrong, and disparages it with his usual vulgarity, he refers to it only very briefly in passing. Engels seems never to have seriously discussed the truth of the theory: it was obviously untrue since it contradicted Marx. His hostile interest was due to the growth of Jevonian socialists—the Fabians, who were thus heretics to be opposed. This, no doubt, is why the first reference is only in 1888, nearly two decades after the theory began.

As to 'scarcity' itself, Professor Robbins gives as sources for his ideas solely writers who wrote after 1870: Menger, Mises, Fetter, Strigl, and Mayer.[2] Even then the idea did not come in with a bang, as had the subjective theory of value itself, and anyone might be excused for not noticing it. Nevertheless the connexion between 'scarcity' and the subjective theory is extremely intimate. For the latter theory, being the first to emphasize demand, was the first to raise the question of the ends of economic activity in any detail. But once we start seriously considering economic ends, whether of the consumer or of the State, we see at once that they compete, and 'scarcity' economics begins.

Now, as we said earlier, Marxist economics had already crystallized by this time. But, what is more important, a system of economics in which all costs are constant with changes in output—for of such a kind must all systems be which base value upon cost alone, be it labour, time, or money cost—is peculiarly ill adapted to any form of scarcity concept. The introduction of such a modification would have required the thorough recasting of the whole structure. For if quantity has no influence upon cost (since cost is constant) nor upon utility (since utility is neglected), the attention of the economist is not likely to be directed towards quanti-

[1] 5 Jan. 1888 to Danielson; 15 Oct. 1888 to Danielson; 8 Feb. 1890 to Sorge; 12 Sept. 1892 to Conrad Schmidt; 29 Sept. 1892 to Kautsky. These letters appear only to be available in Russian translation: *Marx–Engels Sochineniya* (Marx–Engels Institute, Moscow), vols. xviii and xix. This edition, the most complete we have, is notoriously incomplete owing to political vicissitudes.

[2] *An Essay on the Nature and Significance of Economic Science*, 2nd ed., p. 16. There is, however, an earlier reference in H. H. Gossen, *Entwicklung der Gesetze des menschlichen Verkehrs* (Brunswick, 1854), p. 231.

ties of production at all. All the more then will he neglect the question of relative quantities, i.e. the allocation of scarce resources. Decisions as to relative quantities of output are the essence of welfare economics—or of Robbinsian economics, which is the same thing—and the welfare pricing rules are merely tools whereby the correct quantities may be obtained. But in Marxian economics the pricing rules are the essence, for it is through the settlement of prices that exploitation, surplus value, capitalist breakdowns, &c., arise. A system of *Political* Economy, in other words, which analyses the distribution of the national income between classes, and is used as a basis for historical predictions as to the course of *political* events, stands in sharp contrast to a system of *Economic* Economy which is by definition interested in the relative priorities of various kinds of output. To put it another way, Marx did not design his economics to assist in the planning of a socialist society—indeed, he vehemently eschewed all such aims, as he considered them Utopian. His economics set out to prove the inevitability of a proletarian revolution. Hence the very word 'economic' has an entirely different meaning in Marxism—it refers to the 'relations of production', the social or class relations between producers determined by the technique of production used.

It follows that when Marxists speak of the 'law of value', as they very often do, they are not usually thinking of scarcity or social utility. The 'law of value under capitalism' means for them exploitation and the class struggle. When they speak of the 'law of value under socialism' I am bound to confess that I do not see what many of them have in mind at all. In fact they simply talk nonsense, since on the one hand the law of value cannot refer to exploitation, which has been abolished, and on the other the problem of choice is not recognized. Indeed many of them find a rather feeble content for the law of value by referring it to mere accountancy, which does of course persist under 'socialism' (*khozraschot*, sec. 4 below). This is very clearly brought out by the anonymous (and therefore authoritative) article, 'The Teaching of Economics in the Soviet Union',[1] which created such a stir in U.S.A. during the Soviet-American honeymoon. There was a passage (pp. 519–27 in the translation) insisting on the importance of the law of value under socialism, which other Soviet economists had denied. Mr. Carl Landauer[2] was led by this to head his comment 'From Marx to Menger', but inspection of the passage shows that the anonymous author has no notion whatsoever of choice or consumer's sovereignty in mind—indeed has nothing very clear in mind at all. The same must frankly be said of P. M. Sweezy, who in his section 'Law of Value versus Planning Principle'[3] commits the crass error of

[1] *Pod Znamenem Marxisma*, 7–8, 1943, translated in full in *American Economic Review*, Sept. 1944. [2] *A.E.R.*, June 1944. [3] *Theory of Capitalist Development*, p. 52.

maintaining that the one can *replace* the other—as if a rational planning principle could be other than a law of value. Stalin himself makes the same error (sec. 16 below).

In their criticism of capitalism, again, Marxists point—as well they may—to fluctuations in the demand for goods due to speculation or avoidable changes in producers' incomes. But they do not recognize buried among these the fluctuations due to the rational play of consumer's sovereignty. They speak, too, of the 'disproportions' that may arise in an economy—between agriculture and industry in a backward country, between investment and consumption in a stagnant country, or as manifested even under 'socialism' in physical bottlenecks. They seem to think that all is done when these broad disproportions are abolished, not perceiving that the problem of choice exists apart from all disproportions. The trouble is that *there is no Marxist micro-economics*, and therefore no grasp whatsoever of detail. There is only a nineteenth-century macro-economics of the *Stufenlehre* type. Surplus value itself is a macro-economical idea: Marx himself allows that its rate is an average for the whole economy, and that individual prices are settled on quite different principles,[1] and Sweezy abandons all pretence that a Marxist micro-economics so much as exists.[2]

The sole mention of scarcity in Marxist economics is in a somewhat different context, the distinction between 'socialism' (to each according to his labour) and 'communism' (to each according to his need).[3] Marx prophesied that after the proletarian revolution socialism would set in, and a 'dictatorship of the proletariat'. To this, as the economy develops and time passes, succeeds communism, and, with the assimilation of hostile classes, the State—which would in any case be more liberal than the present 'dictatorship of the *bourgeoisie*'—withers away. We pass from the administration of men to the administration of things. Now under socialism there is still 'scarcity', while under communism there is 'plenty'. But the emphasis on the distinction is typically political, not economic. Scarcity raises problems of distribution, class enmity, and administration: 'here is this basket of goods, who shall have how much of it?'; not 'ought it to have been a different basket of goods?' or 'could it have been produced more cheaply?' The *economic* problem of scarcity is still not recognized. Indeed it is a quite misleading truism that for Marxists politics is part of economics. To say the exact reverse is more enlightening. Marxists are politicians, specialists in power. They approach

[1] *Capital*, iii. 12.
[2] Op. cit. pp. 128–30.
[3] K. Marx, *Critique of the Gotha Programme*. The actual names, socialism and communism, are not found here, being a later addition to the canon. Note that Marxian 'communism' is the same as our notion of 'Bliss' (Ramsey, *Economic Journal*, 1928).

all subjects with Lenin's classic question 'Who whom?' In contrast, modern economics starts by asking, 'How much of what?'

3. After Engles (died 1895) this gap in Marxist economic thinking has been extensively documented by marginalist economists[1] until the moment of the first Five-year Plan. Indeed some of these authors asserted that not only central planning but also public ownership was *theoretically* incompatible with rational choice. Incredible as it may now seem, this contention was for long accepted. Soviet economists, behind their iron curtain, stood aloof from the controversy and it was left to Mansheviks, Labourites, Social Democrats, and other unworthy characters to prove the contrary, as for instance in A. P. Lerner's *Economics of Control.* This historical sequence of events cannot have endeared the concept of 'scarcity' to Soviet economists. Indeed the theory of choice is extremely disruptive of all thoroughgoing collectivism. It shows that choices centrally made are more likely to be wrong even than the choices of complete *laissez-faire*, let alone of a *laissez-faire* suitably modified and controlled. Why, for instance, expand or subsidize a nationalized industry when other sorts of commodity are required? It takes us away from the most rational allocation of resources. We shift from Lenin's 'capture of the commanding heights', a purely political argument for nationalization, to—what? It has never been satisfactorily explained, and trustbusting would appear equally satisfactory. Again, all talk of society's common economic goals, of production for use not profit, loses its sting. For it is primarily by profit that we know in any detail what the goals are. We may correct and modify, but we cannot scrap this basis without a vast statistical mechanism for the central decision of preferences such as Gosplan has never dreamed of.

It is the plan era that interests us, however, and we must skip the period 1895–1929, except for the most distinguished of all Marxists active between those years: Lenin. As late as 1917 Lenin demonstrated his total ignorance of the problem:[2]

> With such *economic* prerequisites it is quite possible, immediately, overnight, after the overthrow of the capitalists and bureaucrats, to supersede them in the *control* of production and distribution, in the work of *keeping account* of labour and its products by the armed workers, by the whole of the armed population. (The question of control and accounting must not be confused with the question of the scientifically educated staff of engineers, agronomists and so on. These gentlemen are working today and obey the capitalists; they will work even better tomorrow and obey the armed workers.)
>
> Accounting and control—these are the *principal* things that are necessary for the 'setting up' and correct functioning of the *first phase* of communist society. *All*

[1] F. A. Hayek (ed.), *Collectivist Economic Planning*; B. Brutzkus, *Economic Planning in Soviet Russia*; T. J. B. Hoff, *Economic Calculation in the Socialist Society*, ch. 4 and 5. I am indebted to these authors for many references and quotations.

[2] *State and Revolution*, v. 4. Italics his.

citizens are transformed into the salaried employees of the state, which consists of the armed workers. *All* citizens become employees and workers of a *single* national state 'syndicate'. All that is required is that they should work equally—do their proper share of work—and get paid equally. The accounting and control necessary for this have been so utterly *simplified* by capitalism that they have become the extraordinarily simple operations of checking, recording and issuing receipts, which anyone who can read and write and who knows the first four rules of arithmetic can perform.

4. Note the emphasis on accounting and control. This is still what Communists really mean by economic planning. Much of Soviet economics, especially its monetary theory, places enormous weight on mere accounting. The very word *khozraschot*, which describes all the main rules for managers, means economic *accounting*, not economic *behaviour*: it corresponds more to the English 'economize' than to the German *bewirtschaften*. Yet accountancy—essentially the same under communism and capitalism—with its preference for the average over the margin and for *ex post* over *ex ante*, is peculiarly unsuited as a guide to entrepreneurial behaviour.[1] The principal tenets of *khozraschot* are: minimize costs on your own initiative, avoid loss, and fulfil the plan. When the plan includes the quantities and prices of all outputs and inputs, and even the identities of most of the suppliers and customers, this is very evident from Professor Lerner's Rule,[2] the instruction to managers appropriate in a scarcity-conscious economy. Soviet money, again, is not an allocating but an accounting and controlling (i.e. checking) mechanism. This is principally evident in the role of the banks, which is by no means to allocate credit to its most profitable uses but to check plan fulfilment: no cheque from one factory to another is cleared unless accompanied by an invoice for goods, which the bank checks against its own copy of the plan laid down for the firm.

It follows that the word 'accounting' is very often misused in the literature of rational choice under socialism. It is of course possible for there to be a 'rational system of economic accounting' in a centrally planned society. But this has no bearing on the questions of choice and resource allocation, which are entirely separate, and indeed can very easily be solved without accountancy of any kind, as in the free competition of small peasants. We must not say 'accounting' when we mean 'allocation'.

A reliable and meaningful currency is on a different footing here from mere accountancy. But rational choice would be very difficult in a *Naturalwirtschaft* owing to the extreme difficulty of exchange, and therefore the crippling imperfection of the various barter markets. It follows that those who depreciate the usefulness of money are usually unaware

[1] Cf. Wiles, *Economic Journal*, 1950, pp. 527–8.
[2] A. P. Lerner, *The Economics of Control*, p. 64.

of the problem of choice. And all Communists in fact depreciate money in one way or another. Thus Marx himself, usually so silent about the economics of socialism, says that socialist money is merely a ticket entitling a man to goods, by certifying that he has done so much work:

> In the case of socialised production money-capital is eliminated. Society distributes labour-power and means of production to the different branches of industry. The producers may eventually receive paper cheques, by means of which they withdraw from the social supply of means of consumption a share corresponding to their labour-time. These cheques are not money. They do not circulate.[1]

Money here is neither a store of value, nor a common denominator between various goods: it is not even a unit of account. More extreme Communists than Marx and Lenin have demanded an absolutely moneyless economy.[2]

5. There is thus every reason to expect that the Soviet economy in the Plan era should have no grasp of the problem of choice. Practical evidence of the absence of 'scarcity' is the extreme arbitrariness of *relative* price setting.[3] 'In 1933', says N. Jasny,[4] 'the price of kerosene for technical purposes, f.o.b. destination, was raised about 10-fold, to a level of about 45-fold that of good Donbass boiler coal, f.o.b. mine. In 1949 the same kerosene cost not quite 6 times as much as coal. . . . In 1949 important types of rolled steel cost about 5 to 6 times as much as coal; in the second half of 1950 only about 3 to 1. . . . The official rates for transportation with horses are about 3 times as high as those by truck.' On my own visit to Moscow in 1952 Soviet economists assured me that internal ruble prices did not determine or even help to determine either exports or imports. It did not matter whether the foreign trade corporation made profits or losses in rubles.[5]

The pricing of gas appeared equally arbitrary to Professor Jewkes in Moscow in 1936:

> Apart from an insignificant number . . . all consumers pay a standard rate. . . . The officials were not inclined to attach much importance to multi-part tariffs. Apparently these have not even been considered in Moscow. There seemed to be no great interest in the view that, since the cost of providing gas to some consumers was higher than to others, there might be a case for graduated prices. . . . When I inquired why, with such high profits, the price of gas was not reduced, the answers made it clear that the officials were not accustomed to think in terms of the relationship between the two. Their answer was that they 'thought the price of gas was low enough'. When I asked whether it was considered equitable to gas consumers, as against those who burnt wood in stoves, to make large profits out of

[1] *Capital*, ii. 18. 2.
[2] A. Baykov, *Development of the Soviet Economic System*, pp. 37–38.
[3] I use only more recent evidence. For the earlier chaos cf. Hoff, op. cit., Appendix C. General inflation or deflation are not, of course, relevant here.
[4] *The Soviet Price System*, pp. 9–11.
[5] Wiles, *Soviet Studies*, Oct. 1952. I am indebted to the editors for permission to repeat certain short passages.

them which would go to the revenue side of the local budget, it became quite apparent that I was discussing the question of prices in terms which were strange to the officials. . . . The officials appeared to think that gas and electricity prices were fixed quite independently.[1]

The pricing policy of public utilities is of course quite peculiarly important and complex, since marginal is usually below average cost, and the difference has to be allocated among users—or possibly among taxpayers—in some rational way. Yet Soviet railway rates know nothing of this problem. The accepted proposition that rates should vary between marginal cost for marginal users and 'what the traffic will bear' for intramarginal users is flatly rejected. 'What the traffic will bear' is monopolistic and capitalistic: rates are and must be based on 'cost'—meaning in the context average cost. What this presumably means is marginal cost plus an equiproportional mark-up for all users, but our source is not perfectly clear on what it does mean. Thus the principal problem is not perceived, and the solution preferred—if we can speak of the solution to an unperceived problem—is mistaken.[2]

The arbitrariness of prices in general is stated quite openly by Soviet economists, who regard it as one of the advantages of 'socialism' over 'capitalism', in that it gives an extra degree of freedom.[3] That this freedom is the freedom of the boat without the compass is not, of course, perceived. Voznesensky, at the time (1947) Chief of Gosplan and a Vice-Premier, says:[4]

> The price of a commodity in the socialist society of the U.S.S.R. is based on its value or its cost of production. However, in the interests of strengthening socialism and raising the standard of living of the toilers, the Soviet state itself determines the exact price of every kind of commodity produced in State enterprises or marketed in state trade, and consequently it determines the degree to which the retail price of the commodity deviates from its actual value.

As is well known, until 1949 output value targets were calculated in 'unchangeable 1926–7 prices'; that is, a factory producing miscellaneous small articles would receive a target not in detailed physical quantities but in a global sum of 1926–7 rubles. Now qualities of goods produced in that year—a pre-plan year in which there was little industry—continued to be valued at their then price. But inflation drove up the actual wholesale prices paid about six-fold by 1949.[5] Yet new goods and new qualities of old goods were given notional 1926–7 prices on the basis of their actual cost in the first years of their mass production. This could in

[1] *Moscow in the Making* (ed. Sir E. Simon), pp. 75–76.
[2] Chernomordik, *Voprosy Ekonomiki*, 9, 1948; cf. Dobb, *Soviet Studies*, Apr. 1950.
[3] Wiles, op. cit.
[4] *The Economy of the U.S.S.R. during World War II* (translated by Public Affairs Press, Washington), p. 72. But on balance Voznesensky seems to have favoured the price mechanism: cf. sec. 17 below.
[5] N. Jasny, *Soviet Prices of Producers' Goods*, p. 15.

individual cases be 10 or even 20 times the price of the obsolete quality, so that 'it becomes better business to over-fulfil plans for secondary and over-priced, but actually less expensive categories of output so as to over-fulfil the overall sales plan calculated in 1948 accounting prices'.[1] Probably, as with our own Prices of Goods Act, 1939, unnecessary changes were made deliberately, merely in order to obtain a new price. The distortion of relative prices by this means must have exceeded all distortions due to arbitrary planning. A good example is given by S. V. Marshev.[2] The following table shows the relative prices of a given quantity of petroleum products, the price of light petroleum products being taken = 100.

	1926–7 'escalating' prices	*1948 accounting prices*	*1947 current plan prices*	*1948 new corrected prices*
Petroleum Coke I . .	124	33	11	40
Petroleum Coke II . .	124	33	6	11
Fuel oil	17	16	6	15
Gas	30	2	2	11
Varnish	41	21	7	17
Polymers . . .	1	9	3	15

It should be remembered that the 1926–7 prices quoted are those reigning in 1948, not necessarily those found in 1926–7!

We may draw evidence from Stalin, too:[3]

> . . . the confusion which still reigns in our country in the matter of price policy. Here is one of numerous examples. Some time ago it was decided, in the interests of cotton cultivation, to bring order into the correlation of prices for cotton and grain, to fix a price for the grain sold to cotton growers and to raise the price for cotton delivered to the State. In this connexion our economists and planners submitted a proposal which could not but surprise the members of the Central Committee (sc. of the Party), for by this proposal the price for a ton of grain was about the same as that for a ton of cotton; the price for a ton of grain was thereby made equal to that of a ton of baked bread. To the remarks of the Central Committee members that the price of a ton of baked bread must be higher than that for a ton of grain in view of the added price for milling and baking, that cotton is altogether much dearer than grain, as the prices of cotton and grain *in the world market* testify, the authors of the proposal found nothing sensible to say. Consequently the Central Committee had to take the matter into its own hands, reduce the price for grain and raise the price for cotton. What would have happened if the proposals of these

[1] Marshev, *Neftyanoe Khozyaistvo*, 12, 1948. Reproduced here from Kaser, *Economic Journal*, 1950, p. 89. The '1948 accounting prices' are a modified form of the 1926–7 prices, and evidently open to the same objections.

[2] Loc. cit., quoted Kaser, loc. cit., p. 89.

[3] *Bolshevik*, 18, 1952. English text of B.B.C., summary of world broadcasts. The phrase emphasized (my italics) is curious, as it seems to be an admission that prices are less rational in the Soviet Union than in the 'capitalist' world.

comrades had acquired legal force? We should have ruined the cotton growers and we should have remained without cotton.

6. We turn from price decisions to output decisions. Note, apart from the arbitrariness of price setting—whether right or wrong the price remains arbitrary—the indication in the last sentence that prices do influence quantities produced, at least in agriculture. For relative prices are only a guide to desirable relative outputs. The latter could be achieved in the teeth of wrong relative prices by neglecting the profit criterion.[1] Yet if profits must be sought, relative prices have decisive influence upon relative quantities of output. This is certainly the case in the relatively free and unplanned agricultural sector, as Stalin shows: had the lower price been offered for grain, cotton-growers would have produced less cotton and enjoyed a lower standard of living, just as under 'capitalism'. The possibility of a subsidy, though no doubt present, is not mentioned. The price is not in this instance a *mere* accounting device but also an influence upon output.

The example of cotton is not perfect, since agriculture is not fully planned,[2] not even the compulsory deliveries of industrial crops. That arbitrary prices should lead to arbitrary outputs there is no matter of surprise; the case is not widely different from the effect of arbitrary price controls in a 'capitalist' economy. But even in the planned industrial sector arbitrary prices must have their influence: for although profits need not be maximized, and indeed there is no great personal advantage to the manager in maximizing them, there are a great many output targets expressed in monetary terms only. Such targets can be reached by producing many different combinations of individual qualities. The management tends to choose the easiest of these combinations, often at the expense of the sub-targets for individual qualities expressed in physical terms. The quotation from Marshev in the previous section demonstrates this point, and so do refugees' statements.[3] It appears, then, that in the detailed choice of products, prices do also influence output in the fully planned sector. Thus detailed output decisions in all sectors are hardly less arbitrary than price decisions.

Less serious, but still serious, is the existence of a large class of free goods. I say less serious because if a good is free but scarce it must be carefully allocated between those who demand it; whereas a wrongly priced good may well be allocated according, or partly according, to its

[1] For instance the marginal cost of rail travel from London to Edinburgh is much lower than that of road travel. But for reasons of 'equity' the railway is forced to charge the same mileage rate as on branch lines, which the buses can undercut. Attempts are then made administratively to stop coaches plying between London and Edinburgh.

[2] i.e. it is a sphere in which there reigns 'commodity production' or production for voluntary sale rather than for planned exchange.

[3] Cf. Berliner, *Quarterly Journal of Economics*, 1952, p. 355.

price out of sheer thoughtlessness—and how could a price be wrong in the most advanced economy in the world? Urban ground sites and State loans for long-term investment (i.e. those provided by Prombank and/or from the 'national economy' item in the budget) are free in the Soviet Union. On the other hand, short-term funds (those provided by Gosbank to factories) are, queerly enough, interest bearing, and so of course is the borrowing of the State from the population (small savings). That urban ground sites and long-term loans should be free of price is of course due to Marxist dogma: rent and interest are capitalist categories.

7. Perhaps choices are rationally made when *plans* are drawn up? The published descriptions of this process contain no mention, let alone discussion, of the problem of choice. The formal initiator of the planning process is the Council of Ministers (Cabinet) behind which stands the Party (presumably the Politburo is the organ meant). These choose—on political, economic, and social grounds combined—certain key (*vedushchie*) branches of the economy, and set targets for these.[1] This is the sole point at which rational economic choice is *possible* in the whole process. Proceeding from these orders and from the known existing situation, Gosplan works out suggested requirements for all Ministries, and these make counter-suggestions. The Ministries pass provisional orders down their chains of command (department, trust, factory) and Gosplan passes orders down to the regional and local planning offices, whose function it is to draw up local plans to be co-ordinated into the all-union plan. All these subordinate bodies also produce counter suggestions, and these influence each other and have to be weighed *simultaneously*: there is no question, key branches apart, of deciding one thing and then going on to the next. A sort of 'sense of the meeting' appears to evolve in the course of argument.[2] It is not clear who decides in case of disagreement; Gosplan is officially an inter-ministerial committee and therefore slightly inferior to the Ministries, and the court of appeal is presumably the Council of Ministers. With such a system, and prices as arbitrary as we have seen, political power is clearly decisive in that resultant of forces, the finished plan.

The chief function claimed in Soviet publications for the planning offices is to *reconcile* the targets allocated, by drawing up physical balances of the goods, services, and money concerned. Do the demands and supplies add up to equal quantities? is the question the planners are said to ask

[1] Mirosh nichenko, *Planovoe Khozyaistvo*, 3, 1951; Smekhov, ibid., 4, 1951; G. Sorokin, *Sotsialisticheskoe Planirovanie Narodnogo Khozyaistva S.S.S.R.* (1946), pp. 61 sqq. Cf. *Soviet Studies*, Apr. 1952, pp. 438, 446, and H. Schwartz, *Russia's Soviet Economy*, pp. 156–7, for English summaries.

[2] S. and B. Webb, *Soviet Communism, a New Civilisation?* pp. 626 sqq.; Turetskiy, *Planovoe Khozyaistvo*, 2, 1936. Cf. Schwartz, op. cit., p. 165, for translation. The formation of a French Cabinet would appear to be the nearest Western parallel.

themselves. If not, they suggest alterations. This is the well-known Soviet 'method of balances', which is claimed to be the essence of the matter. But of course, given the possibilities of destocking, foreign trade, and new investment, there are infinitely many alternative plans which possess this rudimentary consistency with themselves. The 'method of balances' is not—nor is it claimed to be—a guide to the best choice. It merely shows whether the choice made is materially feasible.[1] There is a tendency in Western economic circles to suppose that Gosplan has some secret method of choosing between alternatives. 'If only we could get inside that building and see how they *really* do it.' But there is no reason to go behind the published accounts of the process. *Really* Gosplan does not plan at all in the Western sense—the Council of Ministers plans, and Gosplan tells it whether the targets are materially reconcilable. If the literature *says* the 'method of balances' is the essence of the matter, why should it not be telling the truth?

It is not known what criteria the Council of Ministers uses. If it were not merely arbitrary or political, if there were some secret weapon, some Lerneresque Rule, we should quite certainly have heard of it by now. It is impossible that the numerous high-level refugees should not have known, that the Germans should not have captured documents or high Ukrainian officials capable of explaining: exceedingly improbable that the Soviet literature itself should not have boasted of it. The literature we quote below indicates the criteria of choice to be arbitrary, and we have absolutely no reason to doubt this. Indeed our own experience of the British war economy leaves no other conclusion open: planning choices are essentially more arbitrary than competitive ones, and there is no way yet known to statisticians of making them less so.

Of course scarce resources *are* allocated between competing ends—the figures in the plan consist, by definition, of such allocations. But there is no theory as to *how* they should be allocated. Allocation is merely an inevitable by-product of Soviet planning; it is not what it is for. The object of planning is to increase output, or as Soviet economists so constantly put it, 'socialist construction', which since the war has become 'building communism'. It is the ghost of Stakhanov, not Walras, that presides, and we who preoccupy ourselves so much with resource allocation miss the real point of it all. The first question is not 'How much of *A*

[1] Note the similarity to the matrices of W. Leontief, *The Structure of the American Economy*. Soviet planners do not actually use the mathematics of matrices, for there has been no mention of it in their journal, *Planovoe Khozyaistvo*; moreover, the assumption of fixed input-output proportions makes it useless in practice. But for the rest the resemblance is rather striking. The 'key branches' are of course the determinants (only partial) of Gosplan's matrices. It follows, of course, that choice and scarcity problems cannot be solved by Leontief matrices either—nor indeed has anyone pretended that they could. Their uses lie elsewhere. Cf. H. Schwartz, op. cit., pp. 141–9.

must we sacrifice in order to have *B*?' but 'How can we have more of both?'

The advantage of planning in Soviet eyes is that it is a system whereby every citizen is continually under orders to do better than before. It is a method of increasing output. Its highest value is over-fulfilment, i.e. the upsetting of the relative quantities laid down. True, it is insisted that the exact 'assortment' of articles laid down shall be produced in every detail; over-fulfilment in some at the expense of others is greatly frowned upon. But the factory *is* strongly encouraged to over-fulfil the whole plan: i.e. to divert resources that might in the circumstances have been better used elsewhere. To a certain extent over-fulfilment is contradictory to rational allocation: yet nothing is more praised than genuine all-round over-fulfilment.

The technical function of Soviet planners, then, is co-ordination, not choice: a sort of *ex ante* accountancy. If this lesson is fully learned we must give up certain rather slovenly habits of speech and perhaps of thought when discussing centralized planning. The difficulty that there are an almost infinite number of equations to solve arises not once but twice: first in making the right choices and secondly in seeing that they are consistent with each other. The first process is clearly very much more nearly infinite that the second, and it is only the second that Soviet planners attempt. The second only becomes necessary if a plan is made for a certain time ahead, and must be adhered to; if current adjustment of the plan is permitted it disappears. It is thus not a problem in a freely competitive or at any rate a decentralized economy, and this is why it finds no place in our own theoretical discussions of socialism, or in our economic practice. It was, however, a problem in our war economy.

The distinction between co-ordination and choice is blurred by Soviet phraseology. 'Disharmony', 'disproportion', 'incorrect assortment' are commonly on the lips of Soviet economists. They might mean either a violation of the method of balances or an irrational choice of ends. They almost certainly mean the former in nine cases out of ten.

8. We must next examine Soviet economic theory. We can only be quite certain that a Soviet economist is talking about choice when he speaks of the 'distribution of labour between the branches of industry'. He very seldom does. Thus in Lapidus and Ostrovityanov's 538-page *Outline of Political Economy* (Martin Lawrence, 1929: translated) the 'distribution of labour' rates two short passages. P. 12 correctly describes the solution of choice problems by a free market. Pp. 470–5 deal with the solution under planning: they do not give a rational account of it, indeed not an account at all. Cf. the quotation from Stalin in sec. 16 below.

Soviet economists are extremely conscious of the divergence between private (i.e. monetary) and social profit. The literature we examine below bristles with emphatic recognition of this point. This too is a carry-over from before the Revolution. The divergence was then used by socialists (quite rightly) to refute uncritical capitalist acceptance of whatever results the *laissez-faire* price mechanism turned out. It is now used to excuse the absence of any price mechanism at all. From the truth 'you mustn't rely on economic reasons alone, because there are (say) strategic reasons too' has been deduced the falsehood 'you mustn't rely on economic reasons at all'. I had myself several conversations with Soviet economists in April 1952, mostly on or about 'scarcity'. Being unfamiliar with the controversy discussed below, I had perhaps some difficulty in striking a familiar chord. At any rate after exhaustive discussion my colleagues and I could not discover a glimmer of a notion of 'scarcity'.[1]

Controversies between Soviet economists consistently demonstrate that they know nothing of scarcity, or at least that those who come out 'victorious' do not. We must omit for reasons of space the discussions about the 'law of value under socialism' in 1919–21,[2] and the argument between the 'genetical' and the 'teleological' schools of planning in 1926–9.[3] Rich as these are in evidence for our point, it is more interesting to study a recent controversy in which it seems that the scarcity concept has again won a foothold. We refer to the discussion on the choice between investment projects.[4] In this there are many points at issue, of which we shall select two: (1) For the performance of a given task, how do you choose between a variant with heavy initial and low current expenses and one with low initial and heavy current expenses? (2) Should the criterion be the same on all occasions? It is clear that in each case there is a scarce factor whose allocation is to be decided. The Soviet economists agree that in some sense they are talking about the factor time. However, in precisely what sense they disagree—and very naturally, for the precise definition of this scarce factor is abstinence or waiting, a concept most specifically condemned by Marx.[5] Though we must touch on this special difficulty in order to understand the discussion, we are more interested in the light incidentally thrown on the general role of choice in economic theory and in practical planning. Note that an affirmative answer to question (2) means a single price for waiting throughout the economy in all cases where a choice of this kind arises. Soviet economists who give an affirmative answer are seeking to impose a

[1] Cf. Wiles, op. cit.

[2] Brutzkus, op. cit., i. 2–4; Hoff, op. cit., chs. 4, 5; cf. *American Economic Review*, 1944, *passim*.

[3] M. Dobb, *Soviet Economic Development since* 1917, pp. 324–32.

[4] Summarized by Kaplan, *Journal of Political Economy*, Apr. 1952.

[5] *Capital*, i. 22. iii.

restraint upon the government's economic autocracy, and the question is political dynamite, as we shall see. This danger is quite separate from the danger of falling into decadent cosmopolitanism by introducing a rate of interest to answer question (1). Naturally when we consider the arbitrariness of Soviet prices for materials, instanced above, it is evident that the arbitrariness of the price of waiting can do little additional harm in practice. But it happens that arbitrary pricing has been attacked first in this field, and the theoretical reaction is most instructive.

9. T. Khachaturov starts the ball rolling in his *Principles of Railway Economics*.[1] A 'coefficient of relative effectiveness' (CRE) is described $(I_1-I_2)/(C_2-C_1)$, where 1 and 2 are alternative techniques of obtaining a given output for a given number of years. I is the initial expenditure and C the current operating cost *including amortization*. If $I_1 > I_2$ and $C_1 < C_2$ the choice between techniques depends upon the value chosen for the CRE. But this is formally identical with using the 'capitalist' rate of interest (i): for in a non-communist economy the annual cost of each technique would be reckoned at I_1i+C_1 and I_2i+C_2, and the cheaper would be chosen (remember that C includes amortization). Just as a high external rate of interest would incline us towards technique 2, so does a low normative CRE incline them. With the given output and number of years CRE $= 1/i$ or the number of years' purchase.

Of course the author plays this uncomfortable logic down as best he may. Indeed he can urge an essential difference between CRE and i, in that the CRE is only a planner's device without practical financial counterpart. For as we saw in sec. 6 long-term funds are allocated interest free—otherwise this controversy could never have arisen. Anyway, the technique to be chosen is the one for which the internal 1/CRE (*Anglice* the marginal[2] productivity of capital) exceeds or equals the 1/CRE generally laid down for planning (*Anglice* the rate of interest). But this praiseworthy practice has been distorted and neglected by Soviet planners, says Khachaturov:

> Inadequate treatment of this interesting theoretical problem, a rather important one for putting socialist construction into practice, has had undesirable consequences. The very problem of relating capital outlays to operating expenses has often been denied. Instead it has been proposed to decide the question on the basis of a 'total combination' of the advantages and disadvantages—political, economic, military and others—of each alternative, i.e. in fact to abandon calculations and decide the question arbitrarily . . .
>
> Use of the CRE has been found absolutely necessary in the practice of planning, since without it no calculations for comparing investment alternatives can be carried out. Unfortunately application of these methods has been limited to those cases

[1] About pp. 104–18. I have not been able to consult the original but quote the summary translation in Hunter, *Review of Economics and Statistics*, 1950.

[2] Khachaturov explicitly prefers the average to the margin, but some of his supporters just as explicitly insist on the margin. Into this we cannot enter here.

when it would have been completely impossible to get along without them, or when there was fear of making serious mistakes, or seriously underestimating the amount of capital required. In some cases reconstruction or rationalisation projects have not received any systematic economic accounting analysis whatsoever, on the completely mistaken grounds that the concept of profitability is foreign to a socialist economy. . . .

The problem of the level of the CRE has not been concretely worked out in a single one of the studies dealing with the comparison of investment alternatives. In practice its value has been taken as anything from 5% to 20%.

Khachaturov would clearly answer our questions (1) and (2) with a single, notional, rate of interest.

10. The particular problem of scarcity dealt with here is complicated for a Marxist in that, admitting only labour to be a cost, he does not know what it is that is scarce. Khachaturov does not really define what he is talking about, namely abstinence or 'waiting'. It is thus possible for Mstislavsky[1] to dismiss the CRE as 'it is clear to every Marxist that there are in economics no such technical accounting devices, which cannot in their essence possess definite socio-economic reality'. Mstislavsky is, as we shall see, just a Bolshevik McCarthy in pursuit of homeless cosmopolitans. His contributions are of course all the more illuminating for that.

Both Mstislavsky and Khachaturov shed light upon the suggestion of some of our social-democratic planners that prices should only be charged notionally: that they should only be reckoning devices without financial consequences to those 'buying' and 'selling'. This is exactly the status of the CRE, and it clearly has not worked at all. Indeed it is attacked precisely on the grounds of its notional status.

It is the clear necessity of having a time criterion devoid of counter-revolutionary implications that has inspired Strumilin's[2] suggestion of taking the rate of increase of labour productivity with time, and subjecting it to the algebraical treatment normally reserved for the rate of interest. The fascination of this article for Communists and non-Communists alike is that among all its obscurities it contains an original suggestion: if labour is the source of all value and labour productivity rises from year to year, the value of fixed equipment and intermediate materials ('embodied' labour) falls from year to year in comparison with that of 'living' labour quite apart from any wear and tear; and this ought to affect investment policy. Future value = present value $\div (1+p)^t$, where t is time and p the rate of increase in productivity. The

[1] *Voprosy Ekonomiki*, 10, 1948, p. 38. This is the peak period of anti-cosmopolitanism.

[2] *Izvestiya Akademii Nauk, Otdel. Ek. i Prava*, 3, 1946. Cf. Zauberman, *Review of Economic Studies*, xvi. 7–10; *Soviet Studies*, Oct. 1949 (Miller), Apr. 1950 (Eason and Zauberman), July 1950 (Meek and Bettelheim). It is interesting to note that Strumilin has always been haunted by calculation in labour hours; he was a protagonist of using such calculations and doing away with money altogether in the 'law of value under socialism' controversy of 1920–1.

bad logic of this proposal may be considered in a footnote.[1] For our theme it suffices to note that Strumilin is round his first corner: he has answered our question (1) with his impeccably Marxist factor p, which may be used exactly like the rate of interest. But he then wishes to apply a single economy-wide p to all projects, thus answering our question (2) in the affirmative, and by means of some attractive nonsense inserting a little order into the planning process.

It has become *de rigueur* for Soviet economists of all views to consecrate the last paragraphs of their articles to a vague acceptance of Strumilin's proposal. But none of them can really make head or tail of it, and the controversy proceeds as if it had not been written. However, in accepting his factor p the orthodox (e.g. Mstislavsky, op. cit.) insist that there must be a different p for every industry. And logic is of course on their side, for unlike the psychological sacrifice of waiting, which is identical whether the investment is in steel or railways, the rate of increase of labour productivity does differ in the production of different capital goods. Nonsense but rarely gives birth to sense.

11. However, the 'counter-revolutionary' character of suggestions involving interest is merely a complication. Khachaturov's words demonstrate the general absence of any scarcity concept, any theoretical criterion of economic choice. This comes out equally clearly from his opponent. In one of Mstislavsky's other contributions (*Voprosy Ekonomiki*, vii, 1948, p. 132) we read:

> Instead of laying as the foundation of his analysis the real relationships of Soviet economic actuality, as Marxism-Leninism teaches, the author sets out from the unhistorical idea, long unmasked in Marxist literature, of 'maximum effect for minimum outlay'. More than that, Novozhilov[2] in reality reproduces the method

[1] Briefly, (i) p is no substitute for the rate of interest, as it does not deal with the costs of waiting, abstinence, &c., at all; if it has any substance it is an *additional* factor to be considered. (ii) It has no substance. When we compare fixed equipment and intermediate materials with labour we are only comparing indirect with direct labour: e.g. in a steel mill we are comparing the value of bricklayers, constructional engineers, and ore-miners with the value of mill operatives. This comparison is entirely unaffected by the rise *or fall* (how would Strumilin proceed if Soviet productivity fell?) of labour productivity in general. The value ratio of 'embodied' labour (in Marxist terms constant capital, whether fixed or circulating) to 'living' labour (variable capital) is not changed by a general change in the productivity of labour, for the productivity of the labour that from the point of view of any particular works is 'embodied' has risen *pari passu* with the productivity of the labour that is—in that particular works—'living'. From the point of view of the planner, as opposed to that of the particular works, all labour is at any given moment living, including that employed on investment projects. Yet it is the allocation at any given moment of living labour between capital and current production that concerns us. This refutation, moreover, is equally cogent in Marxist or in ordinary economics. Note that the proposal has no connexion, as might be supposed—and is, for instance, by Khachaturov—with obsolescence. The reference is not to the fact that particular better machines render particular worse machines valueless simply through competition, but to the decline in reproduction cost of all machines as labour productivity rises.

[2] Papers of the Leningrad Financial Economic Institute, 3, 1947 (unavailable, it appears, in the West).

of the 'marginal efficiency of capital' used by contemporary vulgar bourgeois economic theoreticians (Marshall, Keynes, etc.). He lays down an 'insufficiency of accumulation', a 'shortage of means of production' (p. 116), just as all bourgeois preachers of 'subjective value' start from the fact of the 'scarcity' or 'insufficiency' of this or that valued good. Following formal logic Prof. Novozhilov reasons thus: 'If it is possible to make more effective outlays than there are means thereto we must concentrate our outlays only on those objects which promise to give a satisfactory effect.' Thence he puts forward the demand for the establishment of a permissible minimum of effectiveness of outlay. . . . Prof. Novozhilov's whole position as it emerges is scholastic and idealistic in its methodology. It breaks away from the substance, the special characteristics and laws of the socialist economy and contradicts it in its content. It is impermissible (*nyelzya*) to connect the effectiveness of capital outlays in the Soviet economy with their influence upon the fund of accumulation. It is a mistake to seek a standard marginal permissible minimum of economic effectiveness for all outlays. The development of the Soviet economy is not subordinate to the assurance of profit, but to the decision of economic-political tasks according to plan, to the raising of the material and cultural level of the toilers. The conditions of each branch of production and of separate objects of construction cannot be unified: witness our many years' practice in socialist construction, which compel us to reject standard norms of the effectiveness of capital outlays.

Novozhilov, then, would clearly answer both our questions with a single notional rate of interest.

12. In his own constructive contribution[1] Mstislavsky answers question (1) by simply adding up the total money costs each year of each variant and choosing the variant that shows the lowest total after an (undefined) 'particular period of time'. In other words he uses no discount for time at all, and shirks the whole problem. To our question (2) the above quotation shows he would return a flat negative: 'the conditions of each branch of production and of separate objects of construction cannot be unified'.

Orthodoxy has bigger guns. Professor Chernomordik's article is considered so important that both German and Chinese Communists have reprinted it.[2]

At present two approaches are employed to solve the problem of effectiveness of capital investment. Some hold that correct, proven evaluation of the effectiveness of variants of capital expenditure can be reached only by applying a co-efficient of effectiveness which plays the same role as interest would. The whole problem, in their opinion is to find—to 'select'—such a co-efficient. Application of such a co-efficient allegedly would remove all difficulties and afford infallible criteria for judgment. This approach is wrong. Others correctly assume that even the best constructed co-efficient in planned socialist society cannot be the only or even one of the major criteria of effectiveness, nor can it be a condition making it possible to solve the complex problems presented by the practice of our construction.

There is no norm or index, though it be reached by the most perfected methods, that can determine what the general volume of accumulation of capital funds in

[1] *Voprosy Ekonomiki*, 6, 1949. Cf. *Soviet Studies*, Apr. 1950, pp. 363–76.

[2] *Voprosy Ekonomiki*, 6, 1949. Cf. *Sowjetwissenschaft* (East Berlin), iii, 1949; *International Economics* (Mukden), Jan. 1950; *Soviet Studies*, Apr. 1950, pp. 331, 359–63. Quotations here from *Current Digest of the Soviet Press*, ii. 1.

the country should be, and where it should be invested—in which branches, in what amounts, in which construction projects and in which regions. Any attempt to depend on a co-efficient of effectiveness would only be a distortion of the principles of socialist planning and, essentially, their negation. . . . Essentially this constitutes rejection of thorough economic analysis of a variant; rejection of careful investigation and weighing of positive and negative aspects of different variants from the point of view of the whole national economy.

Various other criteria for investment are discussed. Nevertheless 'in those cases in which none of the above-mentioned basic economic indices connected with the process of reproduction provides a decisive or significant preponderance in favour of a certain variant, we hold that it is permissible to apply a norm measuring capital investments and operating costs as a kind of secondary, auxiliary index. Of course, such a norm, however secondary and limited in character, cannot be artificial. We must look for this norm in the real conditions of socialist reproduction.' Nevertheless the norm the author goes on to choose is quite simply the rate of interest, in exactly its capitalist form. He suggests no way of deciding on the exact rate of interest, and shirks the question whether it should apply universally.

Although beneath his vigorous language he is surely hedging, Chernomordik appears to answer our question (1) with the rate of interest, and (2) in the negative.

An anonymous review of the scarcity-conscious A. Lurye[1] appears in *Planovoe Khozyaistvo*, iv, 1949 (pp. 85 sqq.). It says:

Propositions which consist in submitting capital construction in the U.S.S.R. to the bourgeois principle of 'efficiency' are in contradiction to the policy of the Party and the State in their work of planning capital investment, and in contradiction to the Soviet method of industrialisation.

The reason given is that the pre-war five-year plans heavily subsidized basic industry: a single economy-wide norm of profitability would have made the whole industrialization of the U.S.S.R. impossible, as it would have diverted funds to light industry. More than one Soviet economist, including Stalin, brings up this objection. Note that it is essentially macro-economic, and misses the whole of Khachaturov's point, which is micro-economic.

13. Not surprisingly, then, we can detect much retraction in Khachaturov's next article.[2] The CRE, even as a criterion between two variants for a given project, has been rightly criticized: 'Such a mechanical view

[1] Like Khachaturov a railwayman: his article was in *Voprosy Ekonomiki Zheleznodorozhnogo Transporta*, 1948. Public utilities seem to provoke economic thought: one is reminded of that morning star of the 'subjectivist' reformation Dupuit, who wrote of the consumer's surplus in the *Annales des Ponts et Chaussées*, 1844.

[2] *Izvestia Akademii Nauk, Otdel. Ek. i Prava*, 4, 1950. Compare *Soviet Studies*, Jan. 1951, pp. 318–22. Khachaturov has also covered his tracks by a loyal and purely liturgical article entitled 'Comrade Stalin and the Railwaymen', ibid. 2, 1950.

rejects the function of the plan, as though the law of value were the determining law of our economy', and so on. However, that interests us less than his most illuminating remarks on the actual process of planning (pp. 243–5):

> The basic criterion of the effectiveness of investment in U.S.S.R. is the speedier building of communism. . . . All these most important planning tasks are established by the directives of the party and the government, the decisions of Party Congress and conferences of the party, the concrete instructions given by our leader and teacher, the great Stalin. Comrade Stalin directly guides the decisions of the most important questions of the development of the national economy of our state, and among them the ēlaboration of the construction programme. . . . The determination of the scope of the projects included in the yearly plan is the most important section of the plan. In accordance with the U.S.S.R. government regulation 'on the system for confirming title lists of construction' (23 Oct. 1934) the decision to build all large—so-called extraordinary (*sverkhlimitnykh*)—projects whatsoever—never mind whether they are classified as all-union, republican or local—is made in every single case by the U.S.S.R. government. The schedule of these projects—their title list—is confirmed by the U.S.S.R. Council of Ministers. . . . Each project can only be begun after the technical blueprint and estimates have been confirmed. The drawing up of all blueprints and estimates is strictly regulated by official technical instructions, conditions, handbooks and price lists.'

We have here an exceedingly clear statement of the political, non-economic nature of Soviet economic decisions. In his next paragraph, which lays emphasis on economic rationality, Khachaturov seems to be recommending rather than describing an actual state of affairs:

> In confirming a project, the considerations taken into account are not only those of the national economy, politics and defence but also cost. The costing and monetary indices of a project are a constituent part of its economic indices in general, which tell us the size of the necessary outlay in labour, materials, fuel equipment and monetary means for the construction and exploitation of the project. Costing and monetary accounts are an extremely important statistical planning-economic index. The significance of costing indices flows from the role of the law of cost in its reformed aspect under socialism: the Soviet government, having mastered the law of cost consciously, uses it as one of the tools of the plan for the goal of fulfilling and over fulfilling it. The government presents the demand that all state resources made available in the plan for capital investments be used in the most effective and economic way, so that *eo ipso* the ruble ensures the most rational choice between the alternative ways of realising the project given in the plan and the most efficient and timely completion of capital works. . . . Cost indices are very important but not final in determining the sphere of rational application of this or that technical improvement: which has much significance in the making of projects. E.g. we know that railway electrification is appropriate with heavy traffic, mountainous country or cheap power. But to answer the question, with *what precise quantity* of freight, with *how steep* a gradient and with *what price* of fuel and electric energy does the transition from steam to electric traction become profitable, we cannot avoid monetary calculations.

Khachaturov is clearly putting up some resistance, even though there follows a passage on the extreme importance of social and defence factors.

14. The confusion that reigns at present, with different CRE's used by different departments, is described by Baumgoltz,[1] who strongly advocates a single CRE, with which he would clearly answer both our questions:

For example in the Ministry of Generating Stations differentiated norms have been established for hydro-electrical projects according to counties (*raions*) and especially according to the local availability of fuel. For constructing hydro-electric stations in the pioneer counties of the Far North, where local fuel is absent, lower values of the CRE are recommended; as a result even extremely large investments in hydro-electricity are recognised as profitable. When the costs of hydro-electricity are compared with the costs of some other way, necessarily very expensive in the given county, of obtaining energy (having either high initial or high operating costs), this already by itself is an argument for building a hydro-electric station, and it would seem unnecessary to lower the CRE for hydro-electricity in addition. And even if it were possible to agree with a special favouritism for hydro-electricity in comparison with other sources of energy, we must reckon with the kick-back, so to speak, of the CRE. For it should be applied in making the project to every element of the hydro-electric station—to the choice of its most profitable size. With a low CRE canals, tunnels, pipe-lines, the cross-sections of cables etc. of 'de luxe' dimensions are selected. There results obvious irrationality: the exploitation of the most difficult regions proceeds with enormous wastage—'not counting the cost'. As we see, the application of differentiated norms is not justified here.

A second example: the transmission of energy by cable is being compared with the carriage of fuel by railway. A generating station can be established in the town and coal brought to it from a distant field, or it can be established in the coal-field and energy transmitted to the town by cable. By railway initial costs are higher and operating costs lower than by cable. Comparison shows the relationship[2] between the saving of operating costs and the excess of initial costs in our example to be 7%. If the indicated decision is to be referred to the Ministry of Communications the cable must be erected, for in the project-making organisations of that Ministry a value of 8 to 12% has been set for the CRE. But if the Ministry of Generating Stations (Hydro and Steam Construction Department) is going to decide, then the cable[3] variant will be recognised as more profitable, as the norms valid in that department have been established between the limits of 4 and 6·6% (15–25 years' purchase). Finally if the question is to be decided by the norms of Teploelectroproyekt MES the railway variant imposes itself still more strongly, for in that large all-union project-making body the principle of minimum costs rules, that is the CRE is assumed to be zero.[4]

Note then that in practice there is no differentiated norm by branches of the economy: it is differentiated by *departments*. The self-same tunnel, projected by the economic norms of railway construction, receives one cross-section, but by the norms of hydro technical construction another. The construction of any equipment, as is known, depends on the CRE. For instance the quantity of iron and copper in electric motors, generators and transformers is decided by comparing the initial cost of the metals with the annual cost of the loss of current in copper and in iron. The relationship depends on the CRE. It would appear that an electric motor of identical technical characteristics, destined for, say electric traction, should have (according to the railway CRE) one set of dimensions, but for hydro-electric stations another; or that for the conditions of the Far North it should be larger, but for regions rich in local fuel smaller. Actually, of course, motors are reckoned by the

[1] *Izvestia Akademii Nauk, Otdel. Ek. i Prava*, 6, 1950, pp. 449–50.
[2] The reader should bear in mind the definition of the CRE in sec. 9.
[3] The author has clearly made a mistake: he means the railway.
[4] This is effectually Mstislavsky's method: sec. 12 above.

economic norms of the producing plant, which have little in common with the norms of other departments.

Supporters of the price mechanism *à tout prix* could hardly wish for a better illustration of what happens when it is neglected. One wonders what Mr. Baumgoltz would make of our own Public Works Loan Board.

15. We may sum the controversy up before coming to Stalin's decision. Few now complain about the CRE itself, as an answer to our question (1). Attention has shifted from the search for a Marxist substitute for it to the advantages of universally applying a single CRE—our question (2). The supporters of the price mechanism, who answer this question affirmatively, are those in touch with reality: the project-makers (*proyektirovshchiky*), who sit in the specialist Ministries and make out detailed estimates for particular schemes. The support for the rate of interest is founded on a practical need. Keynesians, if any be left, who believe that the rate of interest is a purely monetary phenomenon, that Nassau Senior went down the drain when liquidity preference was enthroned, would do well to contemplate this fact. The orthodox, *status quo* party are the ideologues and perhaps Gosplan itself: the journal *Planovoe Khozyaistvo*, whose threatening anonymous review is quoted above, is Gosplan's house organ. In U.S.S.R. 'planners' are quite different from 'project-makers', being occupied not with detail but with the broad sweep, and not with making proposals but with reconciling them. The social utility of a given proposal is not, as we have seen, their business. The comparative silence of *Planovoe Khozyaistvo* in this controversy is thus easily explained: it is simply not concerned. The chief organ of the orthodox is *Voprosy Ekonomiki*, the bell-wether of the economic press: it has published nothing unorthodox. The unorthodox, among whom must on balance be included Strumilin, use the organs of the Department of Law and Economics in the Academy of Sciences, the Leningrad Institute of Finance and Economics, and the technical journals.[1] Very little that is orthodox on this issue has appeared in any of these publications.

The unorthodox have put forward very moderate claims for the price mechanism and for rational choice; they are doubtless ignorant how wide these claims might be. They have asked simply for a single rate of interest with which to choose between investment techniques, given both annual output required and number of years. Indeed the CRE cannot operate unless both are given, and thus its very definition precludes the raising of wider questions. They have raised no claims for consumer's sovereignty and scarcely any that the project maker should be allowed to decide the task of the investment, i.e. determine final output: for of course there are some kinds of final output that a fully fledged rate of

[1] Compare the list in *Soviet Studies*, Oct. 1949, p. 124.

interest, as opposed to the CRE, would render unprofitable by any technique. Stalin (below) is one of the few to realize how all-consuming the claims of rational choice may be, and in warning his opponents of this new ground he warns them off altogether.

16. In 1951 ensues a great silence, except for two orthodox articles in the first number of *Voprosy Ekonomiki*. It is now known that in November 1951 the Central Committee of the Party was discussing the draft of an economic textbook, which deals with the question under review. The question had gone up higher, and no academic would be foolish enough to publish on it. We need not wait for the textbook, for we have the decision it will arrive at on the problem of choice already: in Stalin's recent great pronouncement.[1] The relevant passage is:

> As to the distribution of labour between the branches of production,[2] it will be regulated not by the law of value, which will lose its force by that time (sc. under Communism), but by the growth of the needs of society. This will be a society in which production will be regulated by the needs of society, and the consideration of the needs of society will[3] acquire first-rate importance for the planning organs.
>
> The assertion that under our present economic regime, in the first phase of development of communist society, the law of value regulates the 'proportions' of labour distribution between the various branches of production, is also quite wrong. If this were right it would be hard to understand why light industry is not developed in our country at top speed as the most profitable industry, in preference to heavy industry which is often less profitable and occasionally quite unprofitable. . . .
>
> Clearly, following in the footsteps of these comrades, we should have to give up the priority production of the means of production in favour of the production of the means of consumption. And what does the abandonment of the priority production of the means of production signify? It would mean the destruction of the possibility of uninterrupted growth of our national economy, for it is impossible to effect an uninterrupted growth of the national economy without giving effect at the same time to the priority production of the means of production. These comrades forget that the law of value can operate as the regulator of production only under capitalism, in conditions where the means of production are privately owned, in conditions of competition, anarchy of production and crises of over-production. They forget that with us the sphere of activity of the law of value is restricted by the existence of the public ownership of the means of production, by the operation of the law of planned development of the national economy; consequently it is restricted by our annual and 5 year plans, which are an *approximate* reflection of the requirements of this law.
>
> Some comrades conclude from this that the law of the planned development of the national economy, the planning of the national economy, destroys the principle of profitability. This is quite wrong and quite the contrary is true. If one considers profitability not from the point of view of individual enterprises or branches of production, and not within the span of one year, but from the point of view of the whole national economy and within the span of, say, 10 to 15 years— which would

[1] Op. cit. (my italics). Stalin dates his signature 1 Feb. 1952: presumably the paper circulated unpublished until October.

[2] To repeat: this, if anything, is the Soviet phrase for our 'problem of choice'.

[3] Implying that this is not yet the case. Perhaps Stalin implies that *raison d'état* governs today, while under full Communism the State will wither away and *raison de consommateur* will govern instead.

be the only correct approach—then the temporary and unstable profitability of individual enterprises or branches of production cannot even bear comparison with that higher, stable form of profitability, which the operation of the law of the planned development of the national economy and the planning of the national economy give us, saving us from periodic economic crises which destroy the national economy and inflict colossal material damage on society, and securing for us the uninterrupted growth of the national economy with its high tempos.

In short there can be no doubt that, under our present socialist conditions of production, the law of value cannot be the 'regulator of proportion' in distributing labour between the various branches of production.

Stalin is thus hostile to the general attitude of Baumgoltz, Khachaturov, &c., and answers our question (2) in the negative; but no more than Chernomordik will he pronounce flatly on the status of the CRE for a given single problem of choice between techniques. We may perhaps receive a definitive answer to our question (1) in the textbook; but it is evident that this question is subordinate. We may presume that the whole controversy is now at an end.

17. A final instance from the field of theory. It is significant that one of the alleged reasons for the disgrace (in 1949) of N. A. Voznesensky, Vice-Premier and chief of Gosplan during the war, was his positive attitude to the pricing system. From his celebrated and for a time orthodox work[1] we can extract the following sentiments:

The law of value operates not only in the distribution of goods, but also in the distribution of labour itself between branches of the economy of the U.S.S.R. In this respect, the state plan makes use of the law of value for the proper distribution of social labour between the several branches of the economy, so as to further the interests of socialism. Within the plans for increasing socialist reproduction—and the state economic plans in the U.S.S.R. are such plans—it is necessary to observe certain relationships between industry and agriculture, between the production of the means of production and the production of consumers' goods, between the growth of production and the development of transportation, and between accumulation and consumption. . . . Thus the law of value in socialist economics is a transformed and most elementary law of the cost of production, distribution and exchange of goods, placed in the service of state planning. . . . The conflict between the value of the goods and the rate of profit has been abolished in the U.S.S.R. In a capitalist economy this leads to the situation that whole enterprises, and even entire branches of production, in which the value or cost of production of goods does not yield the prevailing average rate of profit, perish as a result of crises. As is known, the average rate of profit is not a law in the U.S.S.R., and has no bearing on the development of a socialist enterprise or of an entire branch of production. . . . Occasionally, it was necessary during the Patriotic War to violate the laws of development of the Soviet economic system, and to give them a narrowly wartime character. For instance, it is obvious that a multiplicity of retail prices for the same commodities (a situation connected with rationing) is in conflict with the law of value. After the end of the war economy period the need for rationing disappears, the multiplicity of prices becomes detrimental, and it is in turn discarded. The development of normal trade is assured, and the level of commodity prices is, in the final analysis, established to correspond to the value of those commodities.

[1] Op. cit., pp. 87–90.

These passages are embedded, very typically, in a discussion of money, accounting, planning, and 'distribution' in the very different sense of income distribution under socialism and communism. Nevertheless they point on balance to a distinctly Lerneresque viewpoint. The third passage, which appears to contradict the other three, really only means that the law of value operates without violence and friction.

On 24 December 1952 an article in *Izvestia* by the economist Fedosseyev was sharply reproved in *Pravda* by Suslov, then one of the secretaries of the Central Committee of the party. His exalted rank, and the incontinent recantation of Fedosseyev and *Izvestia*, show where official opinion lay. Suslov says, *inter alia*:

> The unMarxist work of N. Voznesensky, 'The Economy of the U.S.S.R. during the Great Patriotic War' was praised to the heavens (sc. in the journal *Bolshevik*, while Fedosseyev edited it). . . . But actually Voznesensky's work only brought confusion into the theory of the political economy of socialism, since it contained a hodge-podge of arbitrary opinions about the role of the plan and the state in Soviet society, and fetishised the law of value as the regulator of the distribution of labour between individual branches of the Soviet economy.

This is the sole objection Suslov raises to the book.

18. To pass from evidence to speculation, another reason why we might expect 'scarcity' economics to be unpopular is the very name. In general, scarcity implies backwardness. Use of the word lays one open to the easy propagandist reply, 'There are no scarcities in the Soviet economy, which is the most advanced in the world.' Useless to object that the scarcity is conceived of only in relation to the number of competing ends, and that those ends can be raised to any quantity by monetary policy, or by a high standard of living; such subtleties will not by-pass entrenched orthodoxy. Secondly and in particular there are some very precise Marxist texts, slogans, and habits of thought which make scarcity an indigestible concept: the revolution is a 'leap from the realm of necessity to the realm of freedom'; by great projects indeed it is possible to re-make nature as well as man; 'communism' is precisely the abolition of scarcity, the rendering to each according to his need, the end of the economic problem. Why study that which will inevitably die? Thirdly, the Bolshevik temperament enjoys exercising power. Superiors are ordinarily rude to their subordinates and expect unquestioning obedience: the technician is often crassly overruled by the party fanatic, in all walks of life. The raising of technical objections to political proposals has cost many heads in Soviet history. Such a temperament is not going to tolerate the introduction of technical restraints hitherto unknown, such as a rate of interest valid for the whole economy. Still less will it respect the relative prices of goods and services generally. Economic organization, like other Soviet organizations, must form an unobstructed channel through which

power may flow from top to bottom. 'Scarcity' considerations would be a blockage in this channel. The foundation of Soviet economics, in theory and practice alike, is not scarcity at all but capital accumulation and technical progress. The theory of choice, apart from the pipe-dream of a centrally calculated welfare function, is essentially devolutionary: it contradicts 'democratic centralism'. 'Canst thou draw out Leviathan with a hook?' It is to be feared that the fine-spun thread of Robbinsian economics will not bear the weight of this animal.

19. This is hardly the place to argue for a similar shift of emphasis in non-Communist economics, though in my opinion we have here much to learn from U.S.S.R. We have not been sufficiently concerned—after Adam Smith—with technical progress. Ricardo's bias was to make economics about the distribution of the national income, and in this Marx on the whole followed him. From Jevons and the Austrians to Keynes the emphasis lay on 'scarcity' and consumer's sovereignty, and since then it has lain on the volume of employment. In the Ricardo–Marx phase technical progress is an agent of social change, if anything harmful to the workers ('artisans and machinery': 'concentration of capital and increasing misery'). In the 'scarcity' phase, which is the most static, technical progress simply disappears from the theory; so that, for instance, the advantage of perfect competition is its adaptability to consumer's sovereignty, and it is not asked whether perhaps monopoly might be more technically progressive. In the employment phase technical progress reappears as a nuisance ('technological unemployment') or an inflationary stimulus ('autonomous investment'). In Soviet economics it occupies its correct place: the place Adam Smith gave it—a cause of the Wealth of Nations. Sheer capital accumulation, divorced from technical progress, has fared but little better at the hands of Western economics; during our employment phase we have valued it chiefly for its multiplier effect, and been if anything afraid of the increase in productivity it brings. The 'warranted rate of growth' is a correct concept for the study of *laissez-faire*. But in a more important sense of the word no rate of growth less than 5 per cent. per annum is warranted. Planning and continuous inflation can float off virtually any increase in production. The Five Year Plans know no warranted rate of growth: to them we may indeed apply Marx's sarcastic description of the capitalism he knew: 'Accumulate, accumulate—that is Moses and all the Prophets.' An economy that adapts itself to internal demand slowly if at all, that substitutes the spur of plan over-fulfilment for the bit of inelastic demand, socialist competition for perfect competition, socialist construction for social security, and forced or directed labour for the free wage bargain—such an economy will raise the production of whatever it does choose to produce more quickly than

ours; *much* more quickly if it can avoid administrative or political mistakes. It is only very recently, under the headings of 'under-developed areas' or 'industrialization and foreign trade', that non-Communist economists have begun to put the emphasis on sheer production.[1] A delicate appreciation of choice problems makes us hesitate before beginning production at all: it is therefore a brake upon production.

20. However, if the non-Communist world has this to learn it remains true that there is in U.S.S.R., despite the timid beginnings of Khachaturov and others, no admission of 'scarcity' to a central place; and therefore no comprehension of the bulk of non-Communist economics, no adequate concept of economic welfare, nor of the international division of labour; even no thoroughly rational theory of what prices and planning are for. The Soviet economy is in great part a choiceless economy. It is run by people who see in economics nothing more than an obsolete nineteenth-century *Stufenlehre*, and has thus developed 'irrationally'. This is not of course to say that in practice engineers, ministers, or others who take the planning decisions pay no attention to 'scarcity'. It must on the contrary be presumed that in a commonsense way they do, just as did British wartime planners. It is to say, however, that they get little help from Gosplan and the economists; and that the choices are very crude indeed, because they get no help from the existing relative prices either. The Soviet politician has not only to modify the economic choice for social or strategic considerations, he has to make the economic choice in the first place. His non-Communist counterpart is better served.

But does the observer in Moscow notice any difference? The choiceless economy solves many of its problems by imitating other economies, and of really obvious deviations from 'rationality' one can name only a few. Small elementary household goods like needles are short while television sets and motor-cars are in good supply. I suspect that canals continue to be built in U.S.S.R. so long after their obsolescence elsewhere simply because there is no proper rate of interest. But it appears that the practical disadvantages of 'choicelessness' have been grossly exaggerated. For notwithstanding all this, timber is exported and bananas are imported; cotton is grown in the south and wheat in the north. The *practical* irrationality of Soviet choices is surely less than the irrationality of the theory.

Indeed, we may go further: Soviet choices are moderately rational in practice, not because of any theory of central planning, but because of actual consumer's sovereignty. The consumer *is* within very narrow limits

[1] The words of Schumpeter and Colin Clark are honourable exceptions. I exclude the 'dynamised' Keynesian economics of Harrod and Hicks: its emphasis is still on full employment *despite* the expansion of productive capacity.

sovereign: he has, that is to say, a veto. He is not planned, he receives no orders as to what to consume. At worst the negative restriction of rationing is placed on him in bad periods. He cannot encourage production by making a certain line profitable, but he can discourage it by making it unprofitable. For unprofitability offends against *khozraschot*—and, we may add, it reduces wholesale and retail turnover below the planned speeds. Variations in the turnover tax cannot altogether prevent this: for an article may be unsaleable even without any tax, and the government is not indifferent to the revenue from the tax. It follows that when an article ceases to sell the plan must be altered. Furthermore, if the decision is one of detail it will have been decentralized in the first place, so that profitability will affect output more quickly.[1]

Another misunderstanding must always be guarded against when it is asserted that Soviet prices and outputs are in actual practice irrational. If we consider sufficiently broad categories the relations are rational enough, *indeed perhaps more rational than elsewhere*.[2] Investment goods are cheap, the turnover tax makes consumption goods dear. Miners get high wages, shop assistants low wages. The skilled and the hard-working prosper, the idle and incompetent suffer. The output of steel rises more swiftly than that of cotton. Nothing could be more sensible, considering the ends the State has set before itself. The irrationality is micro-economic only. Soviet economists only trouble about the larger choices.

In a word, there is substance in the charge that 'scarcity' economics is finicky and academic. Its absence does not prevent the expansion of the Soviet economy. Moreover, the big broad categories such as investment goods and consumption goods are as we have seen correctly priced in relation to each other, perhaps more correctly than under 'capitalism'. The loss of 'welfare' or 'efficiency' through an incorrect micro-economic allocation of resources is surely less than that brought about by unemployment, restrictive labour practices, the refusal to share trade secrets, or the suppression of workable patents, could any of these losses ever be measured. Thus in the Soviet economy there are, as it were, always too few hairbrushes and too many nailbrushes in view of the resources available, while in a 'capitalist' economy this proportion is always more nearly

[1] The following is a very typical statement of the degree to which there is consumer's sovereignty in U.S.S.R.: 'Trade also forwards the development of new branches of industry producing articles of everyday consumption. For instance this is the way in which, under the influence of constantly growing demand, the production of electrical apparatus, refrigerators, wireless sets and bicycles is growing. . . . The trading organizations still study too little the demand of the population and do not struggle hard enough to change the structure of their turnover in correspondence with changes in demand' (*Pravda*, 29 May 1953: these passages could very easily have been paralleled at any time in the Stalin era).

[2] Cf. Kaplan, *Journal of Political Economy*, 1953, pp. 109–10; Bergson, *Quarterly Journal of Economics*, 1950, pp. 408–41.

right. But the production of both these articles is growing at about 10 per cent. per annum in U.S.S.R., and at about 2 per cent. per annum in 'capitalist' countries. In the end the Soviet citizen will be supplied better even with hairbrushes. Professor Hayek's assertion that a centrally *planned* (as opposed to a centrally *owned*) economy cannot rationally allocate resources is proved up to the hilt by Soviet planning practice. But one cannot help wondering, does it really matter?

COMPETITION AND CONTROL IN THE BRITISH SUGAR INDUSTRY

By P. E. HART

1. The existing structure of the British sugar industry is the direct result of government actions in the past and remains subject to legislative control. The Sugar Industry Act, 1942, which was supposed to regulate the trade until the end of the war, was continued under the Emergency Laws (Miscellaneous Provisions) Act, 1947, and is now extended annually by Order in Council. When it is decided to provide legislation on a more permanent basis, the sugar controversy will be revived, and this possibility is one of the reasons for the attempt, made in this paper, to review the organization of this industry and to appraise various sugar policies.

Another reason is the fact that the current international scheme of regulation will shortly be revised. The International Sugar Agreement, 1937, to which Britain and the other important sugar-producing and consuming countries of the world were signatories, was to remain in force until 31 August 1942. However, it was continued throughout the war and is now extended each year to 31 August of the following year. A Committee of the International Sugar Council, appointed in August 1948 to report on the requirements of any new agreement, decided in March 1952 to accept a draft prepared by its chairman as a basis for discussion.[1] This suggests that, by now, discussions on a new agreement are well under way.

Any new agreement on sugar is bound to have important consequences for Britain, traditionally a principal sugar importer and refiner, a beet-sugar producer since World War I, and now also an exporter of refined sugar. The significance of such an agreement is increased by the possibility of a return to the pre-1937 world-wide sugar surplus in the not too distant future.[2]

2. The term 'sugar' embraces a variety of substances, including sucrose, glucose, lactose, fructose, and maltose, but, for our purposes, 'sugar' means sucrose. The commercially important sources of sugar are sugar-cane and sugar-beet. 'In each case, in the presence of chlorophyll, sunlight brings about the combination of carbon dioxide from the air and water to form sugar in the leaf, whence it finds its way to the stem or root for storage as a dilute juice.'[3] In the case of cane-sugar, the juice is

[1] This committee reported to the International Sugar Council in Oct. 1952, and suggested that another International Sugar Conference should be convened to negotiate a new agreement. In Nov. 1952 the Council adopted a resolution requesting the Secretary-General of the United Nations to convene such a conference in London in 1953.

[2] See para. 8, below.

[3] Philip Lyle, 'The Sugar Industry', in *Sources and Nature of the Statistics of the U.K.*, ed. M. G. Kendall.

extracted by crushing the cane: the resultant solution is crystallized to get brown 'raw' sugar. The process of crystallizing the sugar-beet juice is similar, but this juice is extracted by shredding the beet, placing it in warm water, and forming a sugar solution by diffusion. Both raw sugars contain about 97 per cent. sucrose.

In principle, refined sugar is obtained by dissolving washed raw sugar, decolorizing the resultant solution and then recrystallizing it to get 99·8 per cent. sucrose—the white sugar as we know it on the table. These processes involve a loss of weight: a rough and ready relationship to be borne in mind is that 100 tons of raw sugar yield about 95 tons of refined sugar and 5 tons of molasses. The molasses contain about 50 per cent. sugar, and though this cannot be recovered economically as sugar, it can be used for other purposes, for example, in the manufacture of alcohol. After the refining stage, the sugar obtained from cane is almost indistinguishable from that obtained from beet, so they may be regarded as perfect technical substitutes, though their economic substitutability depends on their relative prices.

With this in mind, we may divide Britain's sugar industry into two sections: the manufacturing of raw sugar from beet in the *sugar factories*, and the refining of sugar in *sugar refineries*. This distinction must not be interpreted too strictly; some factories refine the raw sugar they have manufactured and some have even refined imported cane-sugar in the inter-seasonal periods, in order to spread their overheads. Moreover, this classification excludes certain sugar interests in this country, such as the users of molasses (the distillers and the makers of cattle food). It also ignores the sugar market institutions (sugar brokers, &c.) though this probably has more justification in so far as their functions have been taken over by the Ministry of Food. Consequently, for the purposes of this paper, the term 'sugar industry' is simply a shorthand way of referring to the manufacturers and refiners of sugar.

Judged by the numbers employed, the industry is not large. The 'sugar and glucose trades' form the smallest of the 'food, drink, and tobacco trades', with an average annual employment of 18,148 in 1948 and 16,507 in 1935. The number of firms is also small. There are five firms mainly concerned with sugar refining and one firm in the beet-sugar manufacturing section. In so far as the sugar-beet factories use high-cost raw materials, namely sugar-beet, they may be said to have a lower economic efficiency than the sugar refineries, which use the cheaper sugar-cane as a source of raw sugar. Consequently, the sugar factories, but for Exchequer assistance, would make losses, while the refineries can be operated at a profit. This must also be borne in mind when we consider each section of the industry in more detail.

Sugar Manufacturing

3. In this country, raw sugar is manufactured from sugar-beet by the British Sugar Corporation, Ltd. The development of this section of the industry was made possible by State aid, indirectly by protection, and directly by subsidies. It is the story of an infant industry which has never grown up.

In the nineteenth century several efforts were made to establish a beet-sugar industry in the U.K.: small factories were built in 1832 in Essex and in 1868 in Suffolk, but they were unsuccessful. From 1909 onwards, various schemes were put forward for the erection of factories, but none of these came to fruition until 1912, when a factory was constructed by Dutch interests at Cantley, Norfolk. This was operated during the period 1912–15, but made a loss in each year and the original company was wound up in 1916, though the factory itself was reopened in 1920.

Endeavours were made to obtain State assistance, with the result that in 1913 a grant of £11,000 was obtained from the Development Fund in aid of the experiments in the cultivation of sugar-beet. In 1915 the British Sugar Growers' Society, Ltd., was formed and registered under the Industrial and Provident Societies Acts. It proposed the acquisition of a large estate upon which a factory could be built and sufficient beet grown to enable the factory to operate. This suggestion was considered by the Agricultural Policy Sub-Committee of the Reconstruction Committee (the Selborne Committee),[1] which reported in January 1917 and recommended that an experiment on a commercial scale should be made with State Assistance. In 1917 the British Sugar Beet Growers' Society purchased the Kelham Estate, near Newark, Notts., with money advanced by the Treasury from the Development Fund: but the decision to provide money for a factory on the estate was delayed until 1919. In 1920 a company, Home Grown Sugar, Ltd., was floated in order to construct such a factory and to complete the scheme of the British Sugar Beet Growers' Society. The government subscribed half the capital and guaranteed a dividend of 5 per cent. on the remaining half subscribed by the public.

Meanwhile, the owners of the Cantley Factory, which was operated at a

[1] The Selborne Committee was appointed in 1916 'to consider and report upon the methods of effecting an increase in the home-grown food supplies, having regard to the need of such increase in the interest of national security'. The argument that Britain must foster uneconomic production of sugar-beet during peace-time in order to safeguard sugar supplies during time of war commands much support in practice. But it must be remembered that the maldistribution of economic resources implies that a country's national product (and hence war potential) is not as large as it could be. Furthermore, it is possible to keep stocks of commodities to last a short war or until the first sugar-beet harvest. Mr. P. Lamartine Yates in his *Commodity Control* (p. 62) points out that the importing of sugar, as of other commodities, encourages a native merchant shipping industry and nobody would deny that merchant ships are valuable assets in war-time. The strategic argument in favour of economic nationalism has been dealt with by W. Roepke in *International Economic Disintegration* (chapter 8).

loss during the 1920–1 beet campaign, urged the removal of the *excise* duty, which had been placed on home-produced sugar in 1915. The *customs* duty on imported sugar, though abolished in 1874 as part of the Cheap Breakfast Table policy, had been restored in 1901 and by 1918 it was equivalent to 25*s*. 8*d*. per cwt. In 1922 the government removed the excise duty on sugar, thus giving the beet-sugar industry protection of 25*s*. 8*d*. per cwt. During the same year the Cantley factory operated at a profit for the first time. Both factories operated for the 1923 season and the erection of a third factory at Colwick, Notts., was undertaken.

The Budget of 1924 reduced the customs duty (and hence the protective margin of the beet-sugar manufacturers) to 11*s*. 8*d*. per cwt. In July 1924 the government announced its intention of granting a direct subsidy to the U.K. manufacturers of sugar from home-produced beet of 19*s*. 6*d*. per cwt.[1] This decision was given statutory effect by the British Sugar (Subsidy) Act, 1925, which was retrospective to 1924. It was intended to give this direct, but diminishing, assistance for 10 years, by which time it was hoped the infant industry would be mature enough to permit the removal of the subsidy. Though the Chancellor of the Exchequer relied heavily on the infant industry argument to justify this policy, he made it clear that he had other reasons too. He claimed that the subsidies would increase employment, 'revive British agriculture and restore rural England'. However, subsequent events proved him wrong. The rate of assistance fell, and so did the world price of sugar. To save the factories from further financial embarrassment, the British Sugar Industry (Assistance) Act was passed in 1931, offering an advance of 1*s*. 3*d*. per cwt. to all factories on certain conditions; these terms were accepted by twelve factories. The original subsidy, which was to have lasted 10 years, was extended for another year by the British Sugar Act, 1934, pending the report of the Greene Committee, which was appointed in April 1934 to study the industry and to make recommendations on future policy. After the report of this Committee in 1935 (Cmd. 4871), the Sugar Industry (Reorganization) Act, 1936, was passed.

(1) This set up a Sugar Commission charged with keeping under review the growing of sugar-beet, and the manufacture, refining, and consumption of sugar.

(2) It also provided for the amalgamation of fifteen companies, then manufacturing sugar from home-grown beet, into one company, the British Sugar Corporation, Ltd.[2]

[1] The protective margin did not remain at 11*s*. 8*d*. per cwt. because an excise duty was imposed equal to the customs duty on Empire sugar, providing protection against foreign sugar to the extent of the rate of Imperial preference, which was 1*s*. 11*d*. per cwt.

[2] The companies amalgamated were: Anglo-Scottish Beet Sugar Corporation, Ltd.; British Sugar Manufacturers, Ltd.; Central Sugar Co., Ltd.; Ely Beet Sugar Corporation,

(3) In addition, it granted Exchequer assistance in respect of every cwt. of sugar manufactured by the Corporation from home-grown beet up to the standard quantity of 560,000 tons of white sugar per annum. The rate of assistance was to be determined by the Minister of Agriculture after consultation with the Treasury so that there was incentive for 'efficient production' and a 'reasonable rate of profit'.

(4) The Sugar Commission had to keep a register of sugar refiners. The Act also dealt with matters to be provided for in a Sugar Refiners' Agreement and gave the Commission the power to license the refining of sugar.

This Act was still in force when the 1939–45 War began. It was amended by the Sugar Industry Act, 1942, with respect to the war years, and the amended law was continued under the Emergency Laws (Miscellaneous Provisions) Act, 1947. With regard to the beet-sugar industry, the major amendments under the 1942 Act were the change in the method of subsidy from a 'rate of assistance' to 'deficiency payment' and the suspension of the Sugar Commission: the other provisions of the 1936 Act remained and are still in force today.

4. The output and profits of the British Sugar Corporation since its formation are summarized in Tables I and II. Output of home-produced raw sugar has been on the average 26 per cent. of total consumption in the U.K. The accounts of the Corporation show a profit, as in col. 1, Table II, but only because of the direct Exchequer assistance (col. 3). To calculate the real trading balance, we have to deduct this subsidy, i.e. col. 4. This shows the losses of the Corporation and is a measure of its economic inefficiency. The difference between using high-cost raw beet-sugar and low-cost raw cane-sugar is not reflected by differences in refined sugar prices, but by the losses of the Corporation as compared with the profits of the refining firms when they are confronted with the same controlled price of the final product.

This table also suggests that its inefficiency is very high. If col. 1 represents the amount the Corporation should earn—and we know that the Exchequer grants enough assistance to ensure a 'reasonable rate of profit', then its actual losses over the period were almost six times as large as this. Col. 4 indicates the amount to be gained by the community by giving up the beet-sugar industry. In addition, the Exchequer, though not necessarily the community, would gain the difference between the excise duty and the customs duty on some 560,000 tons of raw sugar.[1] Though the

Ltd.; Home Grown Sugar, Ltd.; Ipswich Beet Sugar Factory, Ltd.; King's Lynn Beet Sugar Factory, Ltd.; Lincolnshire Beet Sugar Co., Ltd.; Shropshire Beet Sugar Co., Ltd.; Yorkshire Co., Ltd.; Second Anglo-Scottish Beet Sugar Corporation, Ltd.; Second Lincolnshire Beet Sugar Co., Ltd.; English Beet Sugar Co., Ltd.

[1] The difference between the excise duty and the full customs duty (i.e. non-Empire sugar) is 8*s.* 9*d.* per cwt. So the gain to the Exchequer would be a further £4,900,000 per annum.

TABLE I

Production, Consumption, and Imports of Sugar in the U.K., 1936–50

Metric tons (Raw value)

Period	1 *Home production*	2 *Net imports*	3 *Consumption*	4 *Col. 1 as % of col. 3*
Sept.–Aug.				
1936/7	589,375	2,265,039	2,402,453	24·5
1937/8	426,006	2,435,575	2,344,090	18·2
1938/9	326,868	2,460,478	2,405,041	13·6
Year				
1940	565,612	1,578,794	1,793,000	31·6
1941	472,642	−1,711,032	1,640,000	28·8
1942	534,619	768,451	1,669,000	32·1
1943	629,091	1,506,899	1,715,000	36·7
1944	441,386	1,173,760	1,853,000	23·8
1945	490,745	1,076,070	1,839,000	26·7
1946	529,000	1,485,832	1,871,000	28·3
Sept.–Aug.				
1946/7	628,516	1,758,412	1,999,600	31·5
1947/8	484,496	1,803,898	2,099,000	23·1
1948/9	631,566	1,941,590	2,239,000	28·2
1949/50	522,500	2,322,536	1,986,500	26·3

Source: House of Commons Library. Addendum to Statistical Memorandum No. 15, Table 2, p. 2.

TABLE II

Profits of British Sugar Corporation, Ltd., 1937–50

£

Year	1 *Total disposable income*	2 *Dividend*	3 *Total direct assistance from Exchequer*	4 *Col. 1 minus col. 3*
1937	866,654	150,000	2,561,019	−1,694,365
1938	362,121	163,125	1,108,320	−746,199
1939	31,808	145,000	1,623,680	−1,591,872
1940	854,445	129,375	2,655,319	−1,800,874
1941	379,645	112,500	2,116,063	−1,736,418
1942	284,024	112,500	3,981,512	−3,697,488
1943	254,870	112,500	4,947,095	−4,692,225
1944	276,667	112,500	3,733,030	−3,456,363
1945	261,803	112,500	4,058,756	−3,796,953
1946	294,403	123,750	3,972,226	−3,677,823
1947	544,897	123,750	2,749,780	−2,204,883
1948	393,103	123,750	569,007	−175,904
1949	845,537	123,750	3,441,756	−2,596,219
1950	521,898	118,125	3,847,106	−3,325,208

Source: House of Commons Library. Statistical Memorandum No. 15, Table 14.

gain is small relatively to national income, the elimination of the beet-sugar section of the industry would be an essential part of an economic programme to decrease high-cost production and increase low-cost production—the simplest way of increasing economic efficiency in British industry as a whole.

Sugar Refining

5. We have seen that the development of beet-sugar manufacturing would not have taken place without the help of the State. This is not true of sugar refining. Government aid to this has taken the form of protection from subsidized home production of the beet-sugar manufacturers, who did some refining, and from imports heavily subsidized by governments abroad. However, the problem of 'dumping' troubled the refiners long before the government gave them protection in 1928. Soon after the customs duty on sugar was abolished in 1874, the practice of various Continental powers (Germany, Austria, France, and Russia)—of giving bounties to exporters of beet-sugar in order to sustain home agriculture—began to embarrass British refiners. Refined sugar was imported at a price which left room solely for the lowest cost producers, so that by 1900 there were only 15 refineries left in Britain, whereas there had been 51 in 1864 and 33 in 1882.[1] The home refiners of sugar did not gain relief from foreign competition until the Boer War, when the customs duty on sugar was reimposed, but only to the extent of 4*s*. 2*d*. per cwt. After this war they obtained a more permanent respite as the result of the Brussels Convention, 1902, though the Continental powers did not agree to modify their policies of export bounties until Britain, the major importer, threatened to institute customs duties which would countervail the bounties.

Such was the background to the rise of the most important refiners, Henry Tate and Abram Lyle. A further indication of the high quality of these entrepreneurs is provided by the fact that Bessemer, of steel-converter fame, entered the refining industry and failed in spite of his technical ability. Hutcheson, who did business with Tate, provides a clue to the latter's success.

> His enterprise and clear-headedness were amply indicated when his firm (Messrs. Henry Tate and Sons), in or about the year 1870, secured the 'Boivin Loiseau' process of sugar refining, which it is said, had been previously offered in vain by its inventor to nearly every sugar refiner in Great Britain. At great cost, not only for the patent rights but also for the machinery required to work the process, Mr. Tate thus obtained one chief source of his remarkable success. Again, about the year 1876 the firm of Messrs. Henry Tate and Sons and that of Messrs. David Martineau and Sons, secured the 'Langen' process for making cubes. This process superseded the old sugar loaves and did away with sugar cutting, which, until then, was the

[1] Cf. Noel Deerr, *The History of Sugar*, p. 460.

custom among grocers all over the country, and Tate's cubes have now a world-wide reputation.[1]

Henry Tate had been a grocer and must have known that both the retailers and the consumers would prefer the convenient size of the sugar cube to the cuttings from sugar-loaves.

After the 1914–18 War the number of firms refining sugar was further reduced as a result of amalgamation. In 1921 the family firms of Tate and Lyle combined, and were joined in 1929 by another family with a long tradition of sugar refining, Fairrie and Co.[2] They purchased the refinery of Macfie and Son, Ltd., Liverpool, in 1938 and closed it down exactly 100 years after it had been established. Tate and Lyle, Ltd., also control John Walker and Co. and, together with the Westburn Sugar Refiners, Ltd., control the old Glebe Sugar Refining Co. in Greenock.

6. Before 1928 the refiners of imported raw sugar complained of competition from imported refined sugar, aided by bounties and subsidies from governments overseas, and from white sugar, refined by the beet-sugar factories at home, which were also State assisted. But in that year the government provided protection against imports by retaining the customs duty of 11*s.* 8*d.* per cwt. on imported *refined* sugar and reducing the customs duty on *raw* sugar to correspond to a duty of 9*s.* 4*d.* on fully refined sugar. The home refiners were thus given protection of 2*s.* 4*d.* per cwt. At the same time the government gave fiscal inducement to the sugar factories to concentrate on producing sugar in its raw form.

As a result of this policy, imports of refined sugar were reduced, output of sugar by British refiners increased considerably, the refining companies obtained higher profits and greater security. The refiners undertook to reduce prices by the equivalent of the reduction of 2*s.* 4*d.* in the rate of duty borne by them, and this promise was fulfilled. However, direct competition in the refining of sugar between the refineries and the beet-sugar factories, instead of slackening, became even more intensive. Though the 1928 changes encouraged the factories to concentrate on producing raw sugar, some still found it profitable to refine imported raw sugar during the off-season period.

By 1932, price-reductions resulting from this competition had so reduced

[1] Hutcheson, *Notes on the Sugar Industry*, 1901, pp. 95–96.

[2] Sugar refining is traditionally an industry of family firms. Mr. Geoffrey Fairrie is a director of Tate and Lyle today, but his family entered the industry in 1797 when James Fairrie built the Gartsdyke refinery, Greenock. Similarly, the Martineau family started refining in 1797 and still operate in Whitechapel. Robert Macfie established a refinery at Greenock in 1802, and by 1838 was refining in Leith and Liverpool. Abram Lyle entered the business in 1865 when he became associated with the Glebe refinery. After being a grocer, Henry Tate became a partner in James Wright and Co., refiners in Liverpool: this firm became Henry Tate and Sons in 1870 when the Love Lane refinery was built. The Walker family still owns 45 per cent. of the shares in John Walker and Co., but Tate and Lyle Investments, Ltd., owns the remaining 55 per cent.

the refining margin that some of the factories were in grave danger of being unable to carry on without further assistance; the calculations of the factories, and the rates of subsidy payable, were based on prices which included a 'reasonable' refining margin. To avoid another call on the Exchequer, efforts were made by the Minister of Agriculture to secure some kind of agreement between the two sectors of the industry. As a result, an Industrial Agreement was negotiated in 1933 whereby 500,000 tons of the 1,900,000 tons of refined sugar consumed in Britain per annum, were allocated to the beet-sugar factories as their quota. Some factories did not use all their quotas and sold them to the refiners. The quota arrangements resulted in an increase of 6*d*. per cwt. on the refining margin. The higher prices enabled the factories to carry on without further State assistance and at the same time had a favourable influence on the profits of the refineries. The Greene Committee came to the conclusion that the changes made in 1928 were unduly favourable to the refiners and had placed them in an unduly strong position in negotiating the Agreement.

7. The Sugar Industry (Reorganization) Act, 1936, as we have seen, charged the Sugar Commission with keeping a Register of Refiners, gave it power to license sugar refining, and dealt with matters to be provided for in a Sugar Refiners' Agreement. The Ministry of Agriculture issued the Sugar Refining Agreement (Approval) Order in 1937.[1] This Sugar Refining Agreement allocated the refining of sugar between the following:

(i) The British Sugar Corporation, Ltd.
(ii) Tate and Lyle, Ltd.
(iii) Martineaus, Ltd.
(iv) Macfie and Son, Ltd.
(v) Sankey Sugar Co., Ltd.
(vi) John Walker and Co. (Sugar Refiners), Ltd.
(vii) Westburn Sugar Refiners, Ltd.
(viii) The Glebe Sugar Refining Co., Ltd.
(ix) Merton Grove Co., Ltd.
(x) Manbre and Garton, Ltd.

Parties (ii) to (viii) were referred to as the 'refiners'. The basic tonnage was divided into two main parts: $\frac{5}{19}$ to the Corporation; $\frac{14}{19}$ to the refiners.[2]

The $\frac{14}{19}$ of the basic tonnage allocated to refiners were divided as shown in Table III.

Thus Tate and Lyle, Ltd., was allocated by far the largest portion of the refiners' quota and, if we include quotas of subsidiary companies, it was

[1] Statutory Rules and Orders of the Ministry of Agriculture and Fisheries (1937: No. 221).

[2] Merton Grove Co., Ltd., was allowed 1,050 tons per annum. Manbre and Garton, Ltd., was given the following quotas:
Liquid Sugar—maximum equivalent to 14,970 tons of refined sugar.
Candy Sugar—to customers other than brewers max. 30 tons.
to brewers max. 2,661 tons.

TABLE III

Firm	*Proportion of refiners' quota*
Tate and Lyle	0·74965
Martineaus	0·01229
Macfie*	0·05501
Sankey†	0·05482
Walker*	0·04654
Westburn‡	0·05735
Glebe	0·02434
	1·00000

* Now controlled by Tate and Lyle.
† Now controlled by Manbre and Garton.
‡ Now controlled by Tate and Lyle and Westburn.

responsible for about 87 per cent. of the total output of the refiners. However, we must qualify this figure. In the first place, the allocations were not rigid because quota rights could be bought and sold and, furthermore, we must not forget that the British Sugar Corporation was given a separate quota. After making allowance for this, we may say that Tate and Lyle, Ltd., produce about 60 per cent. of the total refined sugar consumed in this country.

The Sugar Industry Act, 1942, gave the Minister of Food power to modify the Sugar Refining Agreement at the request of the parties concerned. It also suspended the registration of refiners, and transferred certain other functions from the Minister of Agriculture to the Minister of Food. This Act is still effective.

International Sugar Policies

8. At this point we must stress the importance of the connexion between the structure of the sugar industry in this country and the sugar policies of governments overseas. We have already noted the devastating effects of foreign export bounties on the home refineries from 1874 until 1902, when the Brussels Convention eased their position. After the 1914–18 War their security was once again threatened, and they were given protection in 1928.

State intervention in the sugar market was not the exception, but the world-wide rule: even the United States, one of the leading advocates of non-discriminatory international trade, adopted an imperial preference type of agreement on sugar from Cuba and the Philippines. The various government policies to encourage the production of sugar aggravated the general glut during the economic depression in the inter-war years until the International Sugar Agreement, 1937. This was more successful than the Tarafa Agreement, 1927, and the Chadbourne Agreement, 1931, in raising

sugar prices, because it included more countries (both producers and consumers), because it was reached by the respective governments instead of the respective industries, and because it was operating under more favourable conditions—world demand was increasing due to the revival from depression and to stockpiling before the expected war. Britain was a party to this agreement and received, among other things, a production quota of 560,000 tons, the standard quantity under the Sugar Act, 1936. The agreement was still operative when the 1939–45 War began. During this war sugar production in Cuba, the largest pre-war exporter of sugar, increased to record levels, while output in both the Philippines and Java was almost obliterated as a result of war and the Japanese occupation.[1] However, the production of sugar in the Far East will no doubt revive and assume its pre-war proportions. Any future sugar agreement, whether it be on an international or imperial basis, which does not bear this in mind is building on weak foundations. Contraction of sugar production appears to be more difficult than its expansion, so that the revival of Java and the downward inelasticity of the Cuban supply is bound to lead to falling prices.

Nationalization of the sugar industry in this country will not diminish the possibility of low sugar prices in London in the future when, presumably, war-time measures of bulk buying and rationing have gone. The import of cheap sugar is the traditional problem of the British sugar industry[2] and no programme for the economically efficient organization of this industry can be drawn up until some international agreement is reached providing for the concentration of sugar production in those areas where it is cheapest to produce it. Any decision to confine sugar production to those places where it is economic would jeopardize the beet-sugar section in this country. But if non-economic arguments do outweigh the economic case against beet-sugar production—and this is not necessarily true, as we shall see—it is still essential for some decision to be reached on the *level* of uneconomic beet-sugar production which is to be supported on the grounds of military security, welfare of agricultural workers, &c. There is no danger that this level would remove the need for imports of cane-sugar; the standard quantity of 560,000 tons of beet-sugar was about 25 per cent. of pre-war annual consumption, the remaining 75 per cent. was cane-sugar, and, furthermore, even in the worst years of the war we still imported more sugar than we produced at home.[3]

[1] Japan, possessing Formosa, was self-sufficient in sugar, and developed the growth of maize and cotton in the Philippines and East Indies.

[2] This statement refers to both the refining and the manufacturing sections of the industry. The 'hidden hand' cannot be relied upon to concentrate production of sugar in low-cost areas because governments of high-cost areas—in the absence of an appropriate international agreement—will continue to support home sugar production by subsidies, tariffs, &c.

[3] Cf. Table I.

When agreement has been reached on this point, we shall be faced with the problem of allocating the economic production. The most direct method of concentrating production in low-cost areas is to ban export bounties and any other arrangements which shelter uneconomic production. It would no doubt be very difficult to achieve such an agreement; vested interests in the present structure of the world sugar industry, especially the uneconomic high-cost producers, would probably be more powerful than the supporters of the principle of producing sugar as cheaply as possible. On the other hand, the present time is surely more favourable to such an agreement than a period of rapidly falling prices and large surplus stocks of sugar, the situation which existed in the inter-war years and which may return in the not too distant future. Even now, the present rationing of sugar in the U.K. is not due to a world shortage of sugar, but simply to the shortage of dollars; the record production of Cuba has to be paid for in dollars and their scarcity has already led to the accumulation of sugar stocks in Cuba. If an easing of the dollar situation coincides with the revival of Far Eastern output, the fall in sugar prices will be even more rapid and give even more emphasis to the high cost of beet-sugar production. Thus the dollar shortage also facilitates an international agreement as long as it is acute enough to maintain the present levels of non-dollar sugar prices.

As soon as we accept an international policy of this type, we can remove the customs duty on sugar. The case made out by the refiners in 1928 would no longer be valid; under the terms of the international agreement there would be no 'dumping' of sugar and the refineries could be left free to compete with imports of refined sugar. There would be no danger of monopoly and no need for any refining agreement. The problem of determining the output of the refining section and of allocating it between various firms could be left to take care of itself. We can see that nationalization is not strictly relevant to the traditional problem of subsidized imports; it does nothing new to foster economic production of sugar, for this depends on international policy.

Nationalization

9. The brief history of the development of the sugar industry enables us to appraise the 'nationalization-on-the-merits-of-the-industry-concerned' type of argument. The major claims are first, that it is wrong for private shareholders of the manufacturing section to profit from an institution dependent on government financial support, and, second, that nationalization of the refining section would remove the dangers of private monopoly.

The first claim would appear to have some justification. Since its foundation in 1936, the British Sugar Corporation has received about £45

millions in direct assistance from the Exchequer and yet has paid out about £2 millions to shareholders in the form of dividends.[1] On the other hand, the government is one of the larger shareholders, and though it does not own anything like the majority of shares,[2] it does have certain control over policy under powers written into the articles of association of the company. The *prima facie* case for nationalization loses more weight when we relate it to sugar policy.

The manufacture of beet-sugar in this country may be defended on the grounds that in the short run it saves dollars, and in the long run it ensures supplies of sugar during war.[3] If, in spite of these views, it is thought that the growth of sugar-beet in this country is uneconomic and cannot be justified, then there is no point in having the beet-sugar factories of the British Sugar Corporation, Ltd. But if the Corporation is replaced by a government-owned monopoly under nationalization, the vested interests concerned in protecting the existing organization—at the expense of greater economic efficiency—would be even more powerful than they are today. There seems to be a tendency to regard any action performed by a nationalized industry as being in the common interest simply because it is subject to the ultimate control of Parliament, whereas in reality Parliamentary control is no guarantee that a nationalized industry will minimize the opportunity costs of its production. It would be more difficult to abolish the nationalized British Sugar Corporation than the privately owned British Sugar Corporation, because those interested in less costly sugar production can use the injustice argument in support of their major case that the State assistance given to the beet-sugar growers and manufacturers would yield more valuable results if used in developing engineering exports or even in building merchant ships and destroyers to convoy sugar to Britain in time of war. The recognition that subsidizing beet-sugar is neither the most efficient way of saving dollars nor the easiest way to safeguard war-time supplies would imply the abolition of these subsidies, but such a policy could be thwarted more easily by a nationalized board than by a privately owned monopoly. Hence, in view of the possibility that beet-sugar production is or will be too costly,[4] it is imprudent

[1] Cf. Table II. This type of argument ignores the possibility that it may be cheaper to subsidize the B.S.C., Ltd., than to own it. However, in view of the current emphasis on *relative* welfare, it is probable that the case against redistributing income in favour of the shareholders of the B.S.C., Ltd., would outweigh any argument that ownership may be more expensive than subsidization.

[2] 5,000,000 shares of £1 each form the ordinary capital of British Sugar Corporation, Ltd. H.M. Treasury at present owns 750,000. Tate and Lyle Investments, Ltd., owns 726,335 shares.

[3] It also implies a transfer of income from the rest of the community to farmers who grow sugar-beet, and this has been used both to defend and attack the policy.

[4] We are concerned with the undesirability of increasing the power of vested interests in the present form of organization. In order to show that this is undesirable, it is unnecessary

to increase the power of any institution which has vested interests in the present organization of the industry, and the nationalization of the British Sugar Corporation would be doing just this.

10. However, the nationalization of the Corporation does not involve any fundamental change in policy, which, presumably, accounts for the general indifference to this question. The reverse is true in the case of the nationalization of the sugar refiners. Government aid to the refiners has taken the indirect form of trying to repair the defects of a free sugar market rather than to substitute public for private ownership. The socialization of the refiners would be a basic change in the method of approach to the problems of the sugar industry.

The first case for nationalization of the refiners has more weight in practice than in theory. If the British Sugar Corporation is taken over, it may be argued that it would be desirable to nationalize the refiners because 'we cannot afford to nationalize *A* unless we take over *B*; we need the profits of *B* to offset the losses of *A*'. Thus we have a case for the socialization of the refiners similar to the arguments supporting the nationalization of road transport with the railways. The economist may point out that such a policy is uneconomic because it does not minimize the opportunity costs of production. But arguments in favour of reducing costs appear to have been swamped in the past by political expediency, and it may well be that if the British Sugar Corporation is nationalized, the government of the day will be tempted to take over the refineries as well, irrespective of the economic disadvantages of such a policy.

While recognizing that a government may ignore economic principles, we may nevertheless try to appraise an economic argument in favour of nationalization, namely, that it would remove the dangers of a privately owned monopoly. We have seen that Tate and Lyle, Ltd., control about 87 per cent. of the refiners' quota, so there are grounds for saying that it has considerable monopoly power. It is also privately owned. It is true that the Co-operative Wholesale Society, Ltd., is the largest ordinary shareholder in Tate and Lyle,[1] but it holds a very small proportion of the shares, and control is still in the hands of the famous sugar refining families.

to show that the production of sugar-beet is uneconomic at the moment: it is sufficient to show that in the relevant future, if something like pre-war sugar markets return, it will be uneconomic. The Greene report sums up the pre-war position. 'Since, however, on a review of all the facts put before us, we are unable to find positive justification for the expenditure of a sum of several millions per annum on an industry which has no reasonable prospects of ever becoming self-supporting and on the production of a crop which, without that assistance, would at present sugar prices be practically valueless, we cannot recommend the continuance of such assistance' (Greene Report, p. 102, para. 314). (Cmd. 4871, para. 314).

[1] The Co-operative Wholesale Society, Ltd., owns 172,489 of the 7,000,000 ordinary shares issued by Tate and Lyle. The Co-operative Insurance Society owns 29,000.

Of course, the existing government controls preclude any attempt to use this monopoly power to raise prices. Moreover, quite apart from this legacy of war-time regulation, it would still be difficult, if not impossible, for Tate and Lyle to charge monopoly prices for refined sugar. The government-controlled British Sugar Corporation has a quota in addition to the refiner's quota and could compete with Tate and Lyle, Ltd. There is always the danger that imported refined sugar would undermine any attempt on the part of Tate and Lyle to use their monopoly power, and this danger can be increased by further liberalization of international trade. The upshot is that Tate and Lyle, Ltd., may be in a monopoly position but the government can prevent the abuse of such power without resorting to nationalization and without even calling in the Monopolies Commission. It is also worth while to point out that these safeguards would be unlikely to remain if both the manufacturers of raw sugar and the refiners were nationalized; governments appear to be less inclined to attack their own monopolistic actions than to attack those of others. Indeed, the very threat of nationalizing the refiners may be a more effective insurance against monopolistic action than nationalization itself.

11. The conclusions we may draw from this review are first, that the nationalization of the British Sugar Corporation, while increasing social justice in the short run, is likely to make it more difficult to alter the sugar-beet policy, and, second, that the case for the nationalization of the refiners on the grounds of private monopoly is not proven. Neither conclusion implies any approval of the present organization of the sugar industry, but both imply that nationalization would be a move from the frying-pan to the fire. However, it must be emphasized that although nationalization is the major political problem which has confronted the industry since the war, it is of minor importance from the economic point of view. No matter who owns the sugar industry, its future difficulties will be centred on its traditional problem, namely, cheap imports. Nationalization does not contribute anything to the solution of this problem, which is international in origin and must therefore be dealt with at an international rather than a national level.

In the past, international measures have been uneconomic and even self-defeating; they entailed the restriction of the output of cheap cane-sugar and an expansion of high-cost beet-sugar production. This was the result of tariffs, subsidies, quotas bounties, and other controls designed to prevent cane-sugar from competing with beet-sugar. If it is true that some beet-sugar production is necessary for strategic reasons, then completely free competition is impossible, because this would lead to the extinction of the beet-sugar section. But this does not imply that any new international agreement can ignore the competitive solution altogether.

Maximum production from limited economic resources cannot be attained without some mechanism to eliminate the waste generated by restricting low cost in favour of high-cost production. Competition is such a mechanism. This suggests that, apart from any non-economic production of beet-sugar to be permitted on the grounds of national security, the output and prices of beet- and cane-sugar can be decided by the consumers through the price-mechanism.

In short, the approach favoured here is to foster international competition within a broad framework of regulation rather than to hinder it by the development of a miscellany of detailed restrictions and controls which shelter more uneconomic production than can be justified on the grounds of security.

I should like to thank Professors A. Beacham, E. H. Phelps Brown and R. S. Edwards, Mr. S. J. Handfield-Jones, and the management of Tate and Lyle, Ltd., for their invaluable help and criticism of this article in draft.

UNIVERSITY COLLEGE OF WALES,
ABERYSTWYTH.

OXFORD ECONOMIC PAPERS (N.S.)

Volume 5. June 1953. Number 2

CONTENTS

An Inaugural Lecture. By J. R. HICKS 117

Imperfect Knowledge and Economic Efficiency. By G. B. RICHARDSON 136

The Economics of the Schuman Plan. By K. K. F. ZAWADZKI . 157

Production Costs in the Gas Industry. By T. K. GRIBBIN . . 190

Industrial Stoppages and their Economic Significance. By SHEILA V. HOPKINS 209

AMERICAN ECONOMIC REVIEW

CONTENTS

VOLUME XLIII — JUNE 1953

SOME NOTES ON MULTIPLIER THEORY — Ralph Turvey

COMPARABILITY OF LABOR CAPACITIES OF FARM AND NONFARM LABOR — D. G. Johnson

FUTURES TRADING AND HEDGING — Holbrook Working

ECONOMICS OF HIGHER EDUCATION — S. E. Harris

Reviews of Books, Titles of New Books, Periodicals, Notes

The AMERICAN ECONOMIC REVIEW, a quarterly, is the official publication of the American Economic Association and is sent to all members. The annual dues are $6.00. Address editorial communications to Dr. Bernard F. Haley, Editor, AMERICAN ECONOMIC REVIEW, Stanford University, Room 220, Stanford, California; for information concerning other publications and activities of the Association, communicate with the Secretary-Treasurer, Dr. James Washington Bell, American Economic Association, Northwestern University, Evanston, Illinois.

Send for information booklet

THE ECONOMIC JOURNAL

The Quarterly Journal of the Royal Economic Society

JUNE 1953

ACCUMULATION, PRODUCTIVITY AND DISTRIBUTION IN THE BRITISH ECONOMY, 1870–1938 — E. H. Phelps Brown and B. Weber

FOREIGN EXCHANGE RATES AND MONOPOLISTIC COMPETITION — H. Brems

A COMMENT — R. F. Harrod

DEVELOPMENT CHARGES AND THE COMPENSATION-BETTERMENT PROBLEM — R. Turvey

A MODEL OF INCOME DISTRIBUTION — D. G. Champernowne

CHOICE OF PSYCHOLOGY AND AS ECONOMIC ASSUMPTION — A. L. Macfie

NOTES ON USER COST — A. D. Scott

WAGE CLAIMS AND THE PACE OF INFLATION (1948–51) — A. T. Peacock and W. J. L. Ryan

Reviews, Recent Periodicals, and New Books

MACMILLAN & CO., LTD., LONDON, W.C.2

Application for Fellowship to

Royal Economic Society, 4 Portugal Street, London, W.C.2

Annual Subscription £1 10 0 — Life Composition £22 10 0

ECONOMICA

Published Quarterly

ECONOMICA is published quarterly by the London School of Economics and Political Science in February, May, August, and November, and is devoted to research in the fields of Economics, Economic History and Statistics. The Journal is under the direction of an editorial board composed of Professor R. G. D. Allen, Professor T. S. Ashton (Acting Editor), Professor W. T. Baxter, Professor E. H. Phelps Brown, Sir A. M. Carr-Saunders, Professor R. S. Edwards, Professor J. E. Meade, Professor F. W. Paish, Professor Sir Arnold Plant, Professor Lionel Robbins, Professor R. S. Sayers (Acting Editor), with Mr. A. T. Peacock as Assistant Editor.

33rd Year — New Series — Volume XX — No. 79 — August 1953

CONTENTS

Monopoly and Economic Progress — *J. Jewkes*

Optimal Production with Fixed Profits — *Marcus Fleming*

The Element of Lottery in British Government Bonds, 1694–1919 — *Jacob Cohen*

A New Method of Relating British Capital Ownership and Estate Duty Liability to Income Groups — *A. M. Carter*

A Refutation of Mr. Bernardelli — *Kelvin Lancaster*

Ricardo's Correspondence — *T. W. Hutchison*

BOOK REVIEWS — BOOKS RECEIVED

The price of ECONOMICA (New Series) is 8*s.* 6*d.* a number, or £1. 5*s.* per annum post free. Copies may be obtained on order from the Publications Department of the London School of Economics, Houghton Street, Aldwych, London, W.C. 2, to which address all subscriptions and business communications should be sent.

PRINTED IN GREAT BRITAIN

AT THE UNIVERSITY PRESS, OXFORD

BY CHARLES BATEY, PRINTER TO THE UNIVERSITY